Network Enterprise Technology Command (NETCOM)

Final Presentation

05/03/2022

<u>Team:</u> Lydia Barit, Katy Dula, Blake Jacobs, Harrison Leinweber, John McCormick, and Roberts Nelson

Heinz College of Information Systems and Public Policy
Carnegie Mellon University

Carnegie Mellon University
Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Agenda

1. Team Introductions
2. Background, Problem Statement, and Purpose
3. Current Workflow and Solution
4. Project Components
   a. xStream Introduction
   b. Data Preparation
   c. Model Evaluation
   d. Post-hoc Explanation
   e. Demo: User Interface
   f. Risk Response Guidelines
5. Conclusion
6. Future Work & Acknowledgements
7. Q&A

Model Evaluation

Katy Dula

MISM-BIDA

(Process Organizer)

Harrison Leinweber

MISM-BIDA

(Cyber Security Analyst)

John McCormick

MISM-BIDA

(Data Scientist)

Post-hoc Explanations

Blake Jacobs

MISM-BIDA

(Chief Systems Administrator)

Bobby Nelson

MISM-BIDA

(Financial Manager)

Risk Guidelines

Lydia Barit

MSISPM

(Project Manager)

Carnegie Mellon University
HeinzCollege
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Background and Problem Statement

NETCOM → ***Mission:*** *"ensure freedom of action in cyberspace while denying the same to our adversaries"*

Data Science Directorate (DSD) → *Contribute to this mission in a* ***data-centric way***

DSC-PIT → *Provide* ***data science-based analysis*** *to inform decision making*

Problem statement: NETCOM DSD needs a way to automate the process of anomaly detection and evaluate the prospective models to do so

Purpose

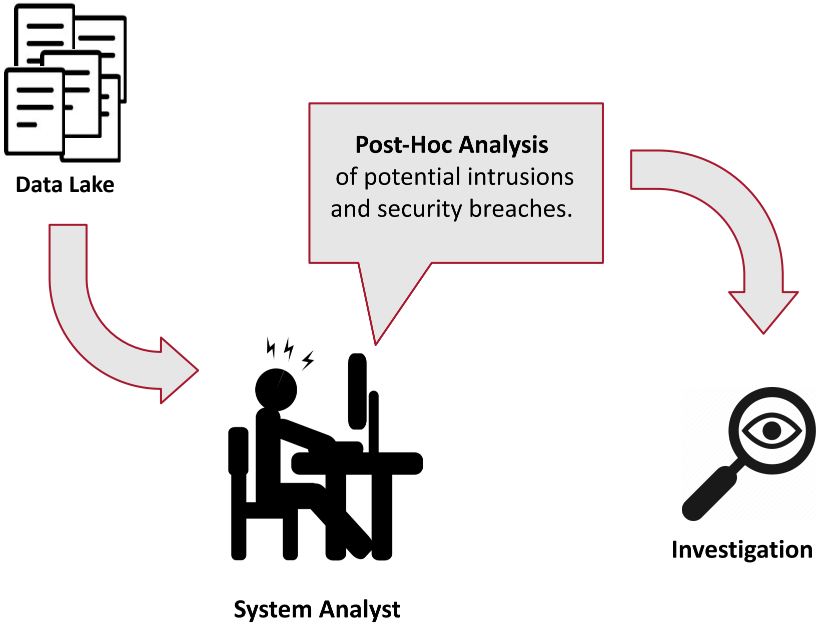
The purpose of our project is to help NETCOM understand the **effectiveness of anomaly detection models** in use cases pertinent to the Command's mission and effectively **incorporate these models** into its current workflow.

DSC-PIT Strategic Capabilities

- **Predictive Analytics/ML**
- **Academic Partnerships**
- **Cyber Security Analytics Products**

xStream
Dr. Leman Akoglu

User Story: Current Workflow



- Current Post-Hoc Analysis does not find many meaningful anomalies
  - Inefficient investigative efforts
  - Limited means to rank or prioritize observations
- Improved algorithms may not be human interpretable
  - Can analysts trust?
  - Allow for further implementation?

Objectives and Impact

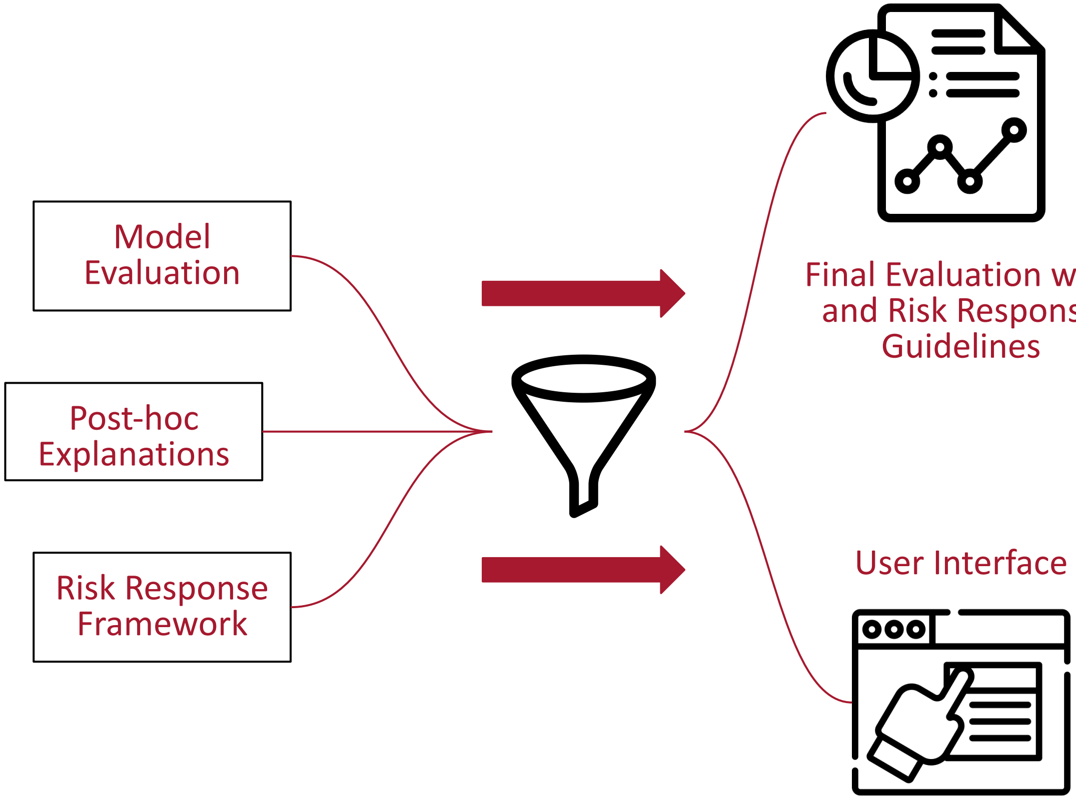
Objectives

- Streamline anomaly detection and investigation
- Provide an understanding of how xStream compares to other models in particular use cases
- Equip NETCOM DSC-PIT with relevant tools to make subsequent risk-based decisions

Mission Impact

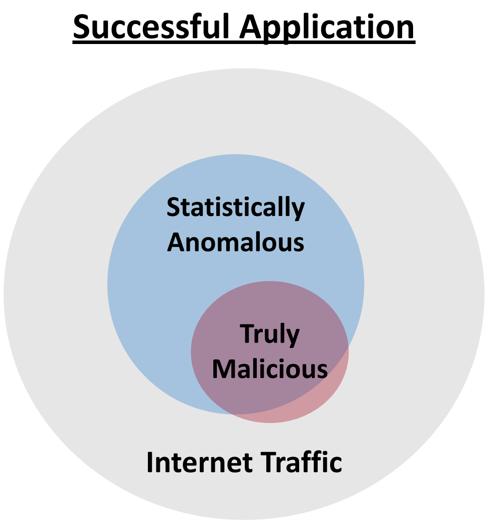
- Our work will allow Soldiers to triage potential threats and vulnerabilities and focus investigative manpower on the most critical cases
  - Increase investigator's precision while maintaining recall

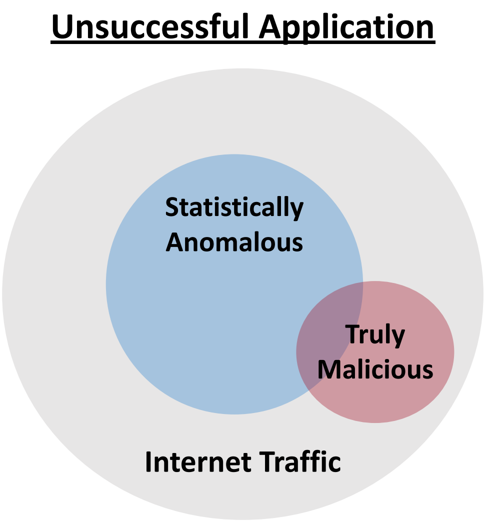
Solution


Model Evaluation
Post-hoc Explanations
Risk Response Framework
Final Evaluation with and Risk Response Guidelines
User Interface

Anomaly Detection for Network Security Systems

- Different tasks that may not necessarily overlap
- Often threat actors intentionally make malicious traffic/domains appear benign
- Successful application depends both on the data-set and the anomaly detection algorithm





Algorithm Overview

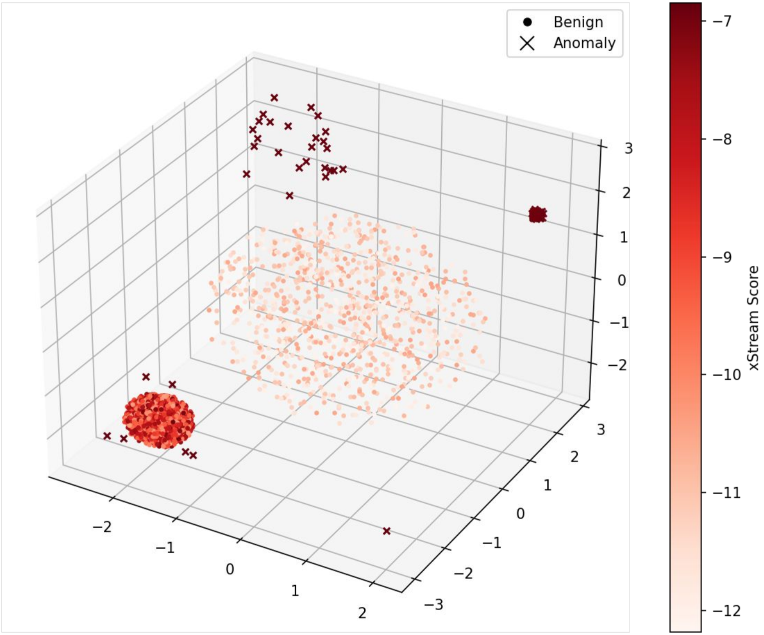
Anomaly Detection Algorithms Considered:

- xStream
- Isolation Forest (iForest)
- Local Outlier Factor (LOF)
- One-Class SVM (OCSVM)

Algorithm Overview

Anomaly Detection Algorithms Considered:

- xStream
- Isolation Forest (iForest)
- Local Outlier Factor (LOF)
- One-Class SVM (OCSVM)



Emaad Manzoor, Hemank Lamba, Leman Akoglu. Outlier Detection in Feature-Evolving Data Streams. In *24th ACM SIGKDD International Conference on Knowledge Discovery and Data mining (KDD).* 2018.

Algorithm Overview

Anomaly Detection Algorithms Considered:

- xStream
- Isolation Forest (iForest)
- Local Outlier Factor (LOF)
- One-Class SVM (OCSVM)

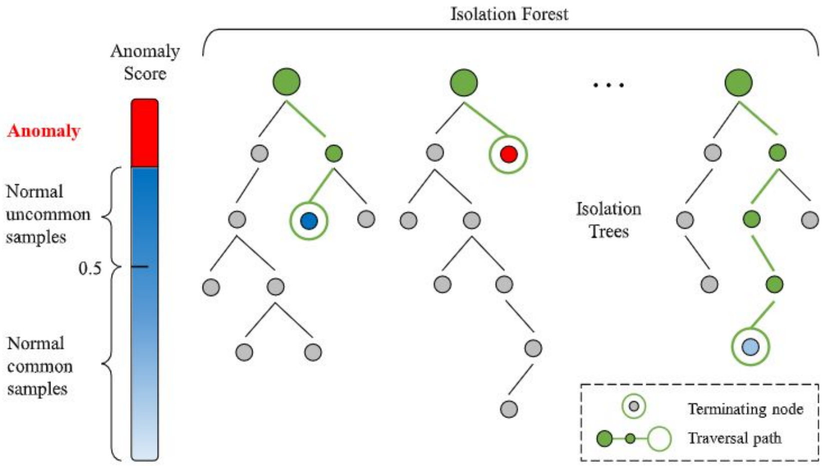


Fig. 3. Anomaly detection with *i*Forest.

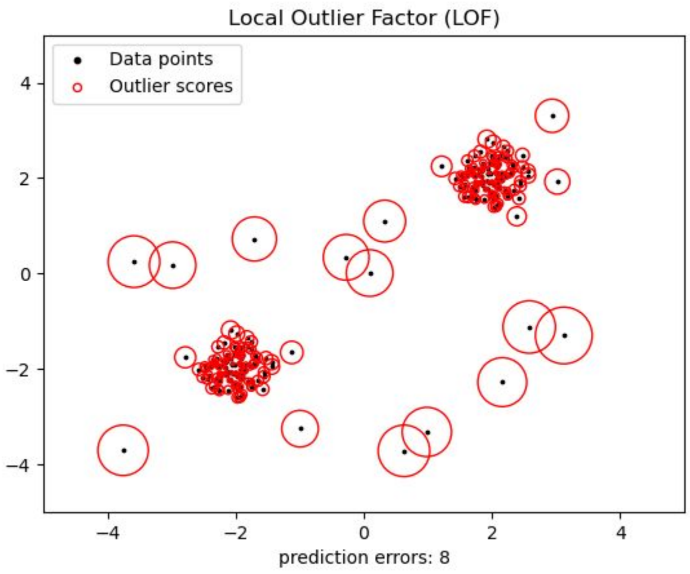
Liu, Fei Tony, Kai Ming Ting, and Zhi-Hua Zhou. "Isolation forest." *2008 eighth ieee international conference on data mining*. IEEE, 2008.

Chen, Hansi, et al. "Anomaly detection and critical attributes identification for products with multiple operating conditions based on isolation forest." *Advanced Engineering Informatics* 46 (2020): 101139.

Algorithm Overview

Anomaly Detection Algorithms Considered:

- xStream
- Isolation Forest (iForest)
- Local Outlier Factor (LOF)
- One-Class SVM (OCSVM)

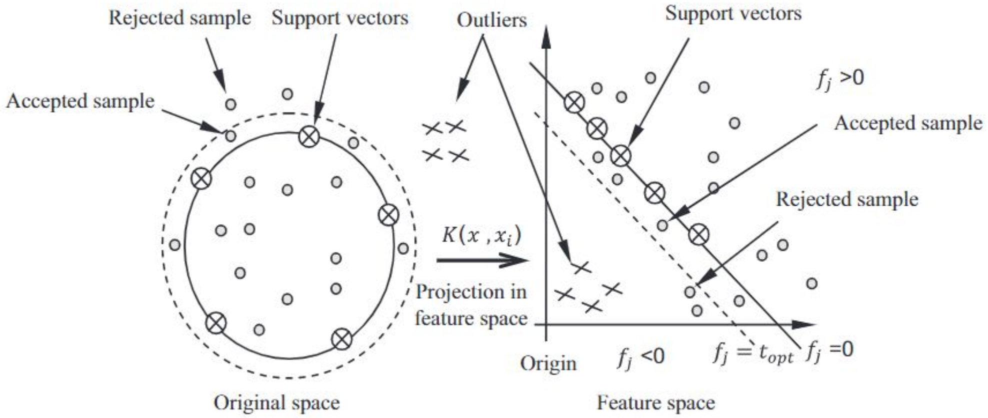


https://scikit-learn.org/stable/auto_examples/neighbors/plot_lof_outlier_detection.html

Algorithm Overview

Anomaly Detection Algorithms Considered:

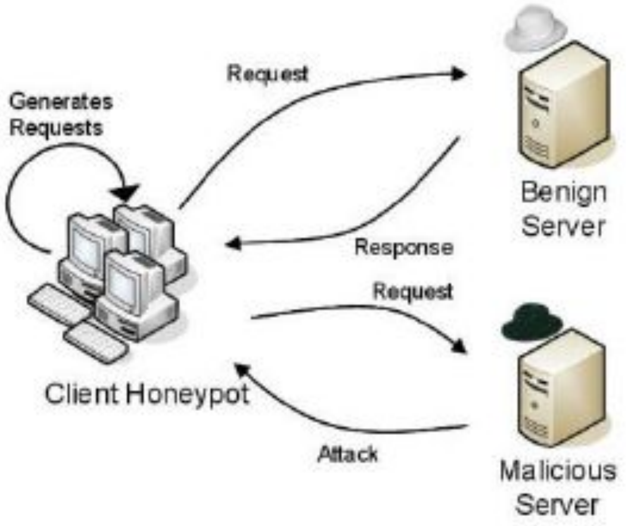
- xStream
- Isolation Forest (iForest)
- Local Outlier Factor (LOF)
- One-Class SVM (OCSVM)



Guerbai, Yasmine, Youcef Chibani, and Bilal Hadjadji. "The effective use of the one-class SVM classifier for handwritten signature verification based on writer-independent parameters." *Pattern Recognition* 48.1 (2015): 103-113.

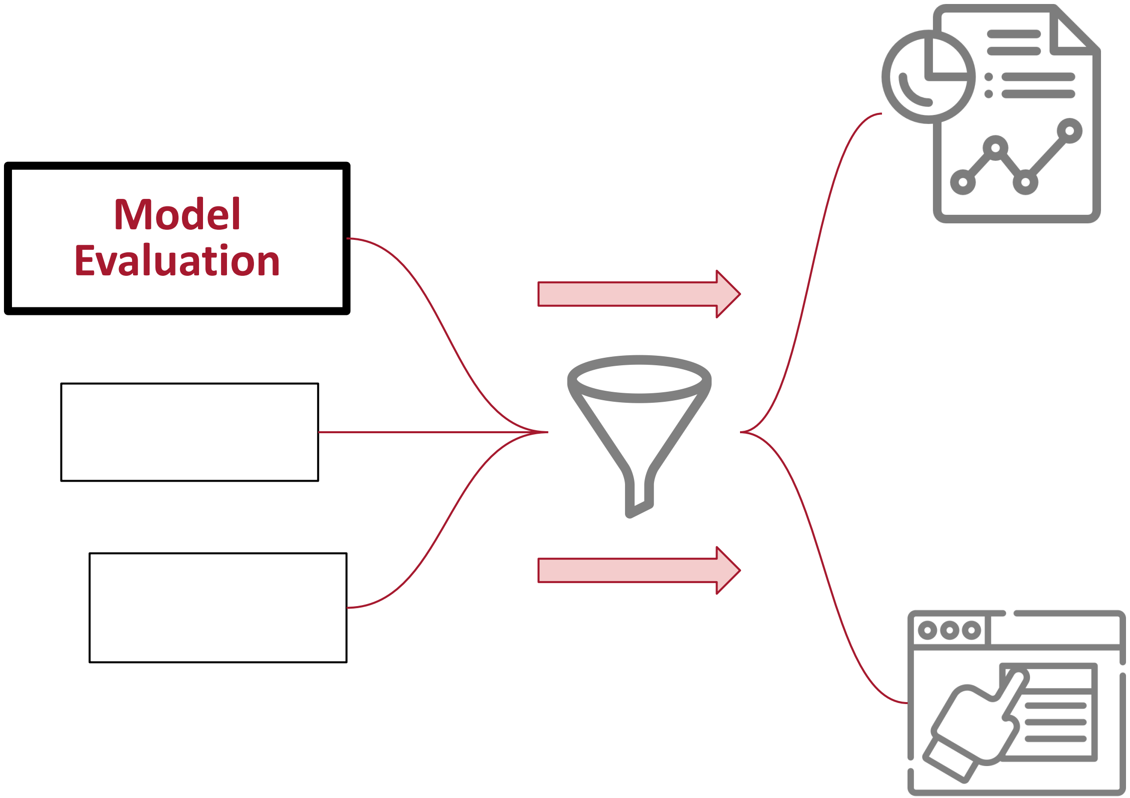
Introduction to Kaggle Dataset

Client Honeypot Technique:



Example of Features:

- Header Content Length
- URL length
- Registration Date
- DNS Packet Count


Model
Evaluation

Model Evaluation: Use Cases

1. Cyber Intrusion Web Attacks
   a. CIC-IDS 2017

2. Malicious URLs
   a. Kaggle Malicious and Benign Websites
   b. NETCOM enriched Dataset

Use Case One:
Cyber Intrusion Web Attacks

Introduction to CIC-IDS 2017 Dataset

Cyber Intrusion Attacks:

- DoS - GoldenEye
- Heartbleed
- DoS - Hulk
- DoS - Slowhttp
- DoS - Slowloris
- SSH - Patator
- FTP - Patator
- Web Attack - Brute Force
- Web Attack - XSS
- Infiltration
- Botnet
- PortScan
- DDoS

Example of Features:

- Packet Length
- Flow Duration
- Various Packet Flags
- Bytes sent in initial window
- Destination Port Number

Preparing the Data

CIC-IDS-2017 Dataset

- ❖ 79 features
- ❖ 2,830,743 observations

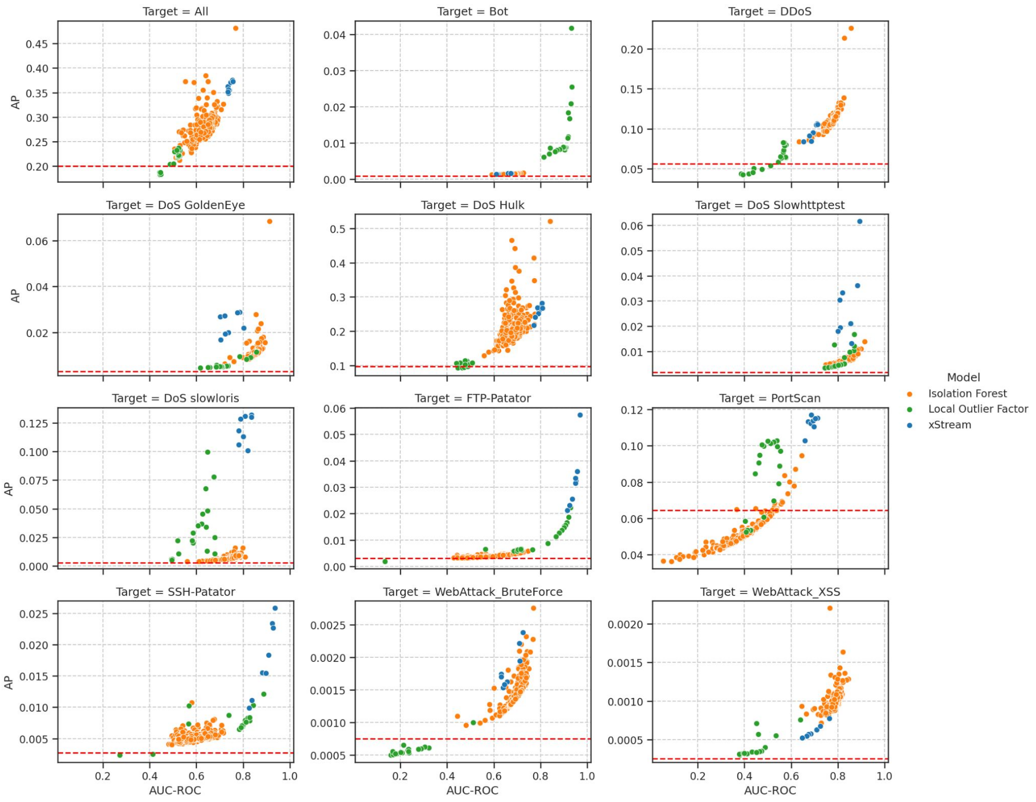
Data Preparation Process

- ➔ 2867 incomplete observations dropped
- ➔ 151 OHE port number features added
- ➔ 8 empty features dropped
- ➔ 59 features normalized

Clean Dataset

- ❖ 230 features
- ❖ 2,827,876 observations

Model Evaluation: Performance by Attack Type

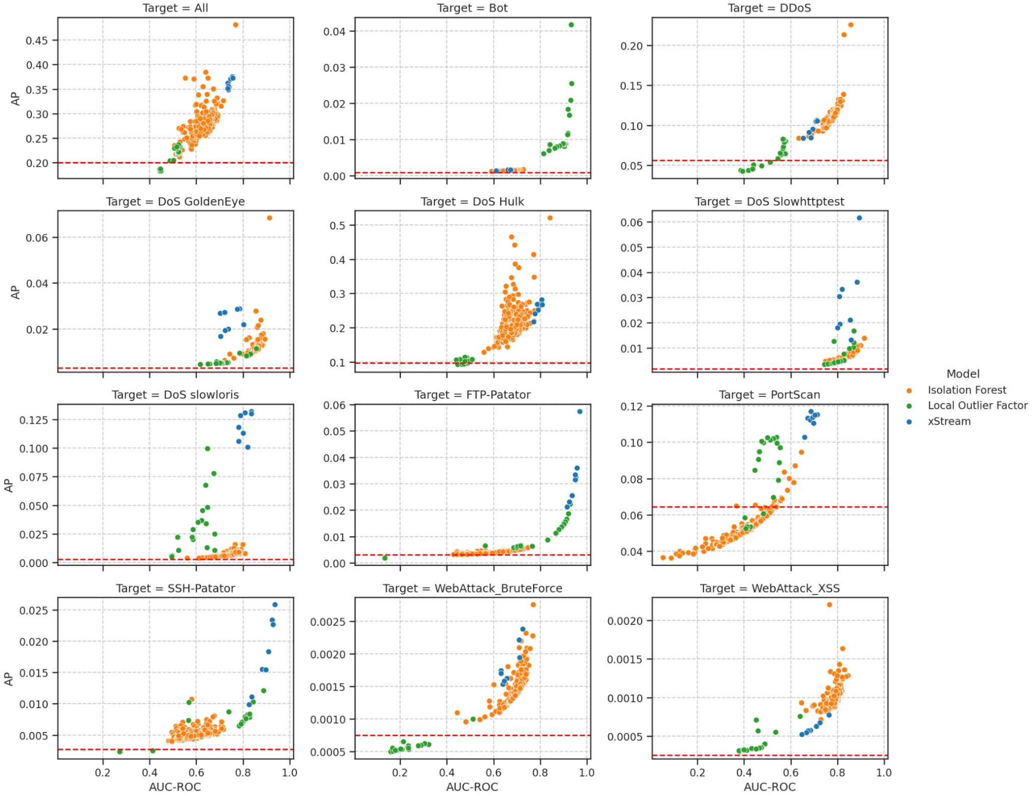


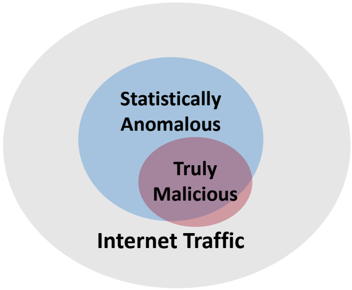
- xStream is overall the better instrument
- xStream does worse in certain attacks:
  - Bot
  - DDoS
  - WebAttack XSS

Isolation Forest xStream Local Outlier Factor

Carnegie Mellon University
Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Model Evaluation: Performance by Attack Type





Anomaly Detection seems to work well on this dataset

Isolation Forest xStream Local Outlier Factor

Use Case Two:
Malicious URLs

Preparing the Data

Kaggle Dataset

- ❖ 20 features
- ❖ 1,781 observations

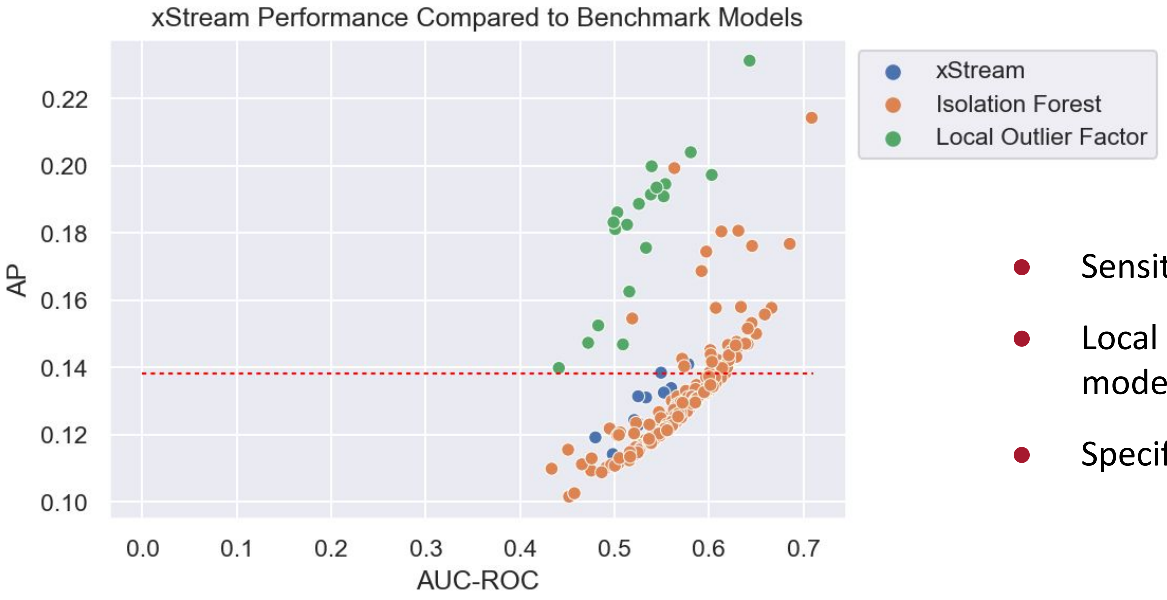
Data Preparation Process

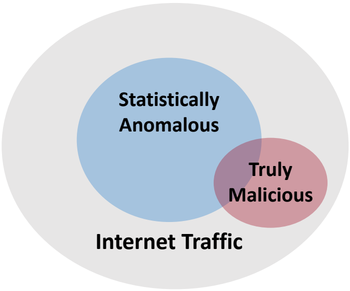
- ➔ 3 features dropped
- ➔ 5 features OHE
- ➔ 6 features with null values cleaned
- ➔ 13 features normalized

Clean Dataset

- ❖ 455 features
- ❖ 1,781 observations

Model Performance

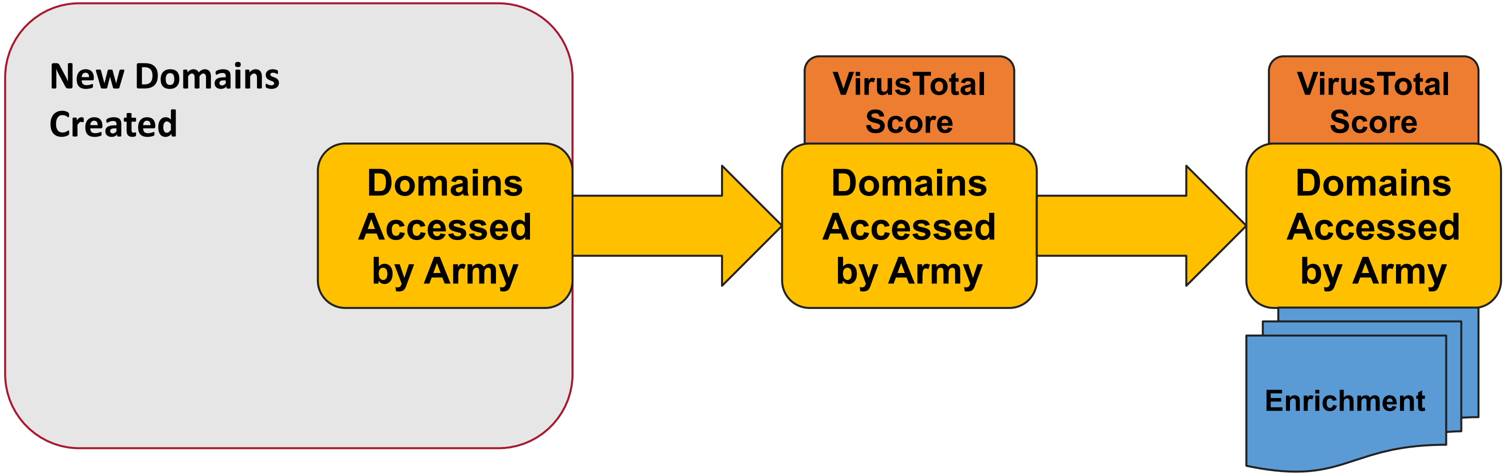




- Sensitive to hyperparameter tuning
- Local Outlier Factor outperforms the other models
- Specific realizations of iForest perform best

Isolation Forest xStream Local Outlier Factor

Introduction to NETCOM Enriched Dataset



Carnegie Mellon University
Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Preparing the Data

NETCOM Enriched Dataset

- ❖ 25 features
- ❖ 696 observations

Data Preparation Process

- ➔ 12 features dropped
- ➔ 3 features OHE
- ➔ 2 features binarized
- ➔ 12 features normalized
- ➔ 2 features cleaned with dictionaries
- ➔ 1 feature engineered
- ➔ 1 feature imputed missing values

Clean Dataset

- ❖ 606 features
- ❖ 696 observations

Proxy Target Variable

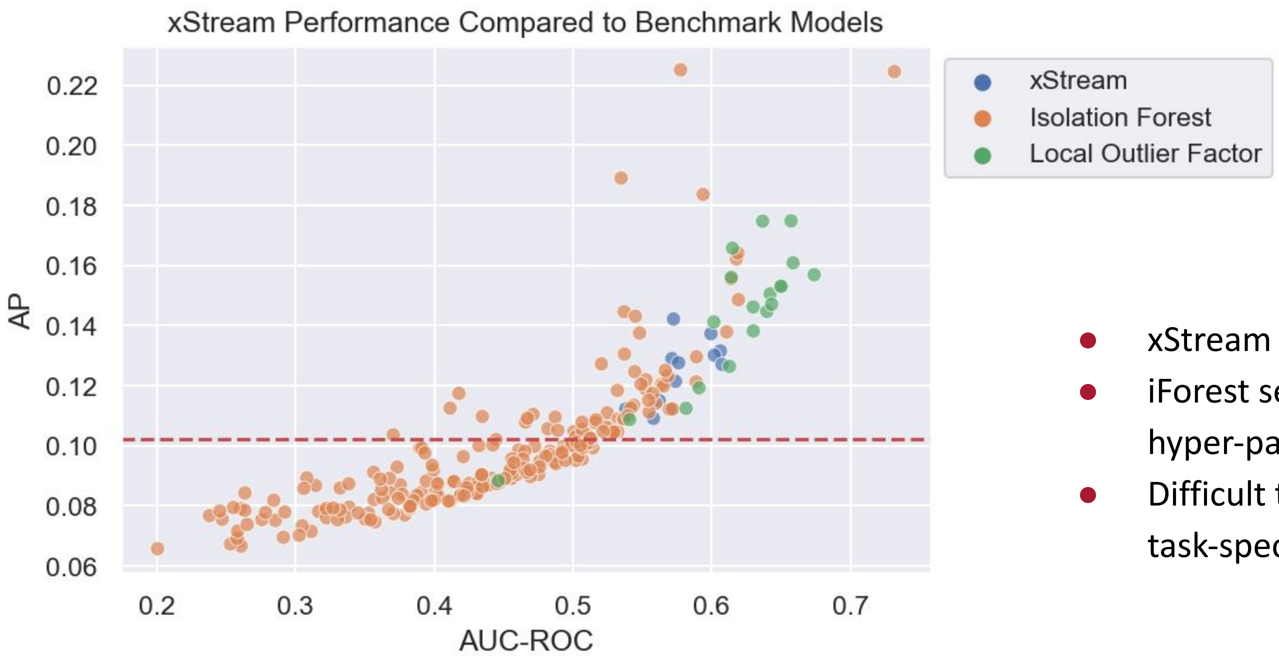
VT_SCORE	
0.0	625
1.0	25
2.0	11
3.0	8
4.0	8
5.0	5
6.0	3
7.0	3
8.0	3
9.0	2
10.0	2
11.0	1

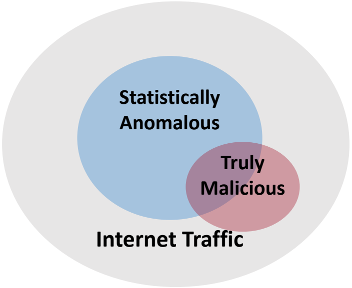
Assumes that the VirusTotal Score Accurately Detects the Malicious sites that NETCOM is interested in

"Benign"

"Malicious"

Model Performance

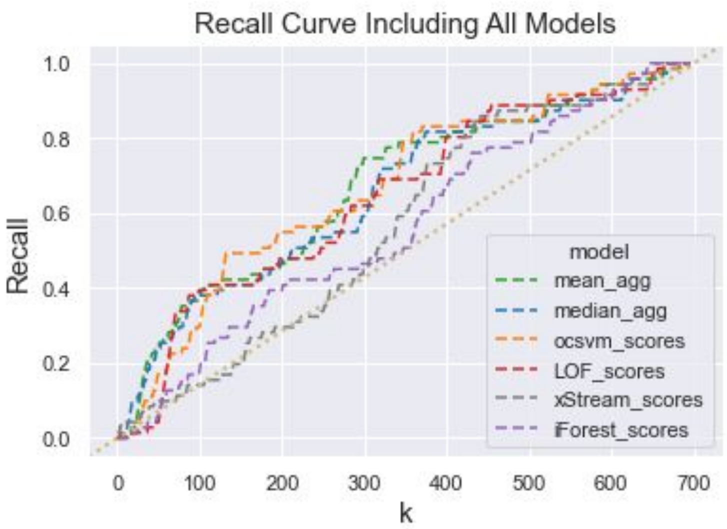


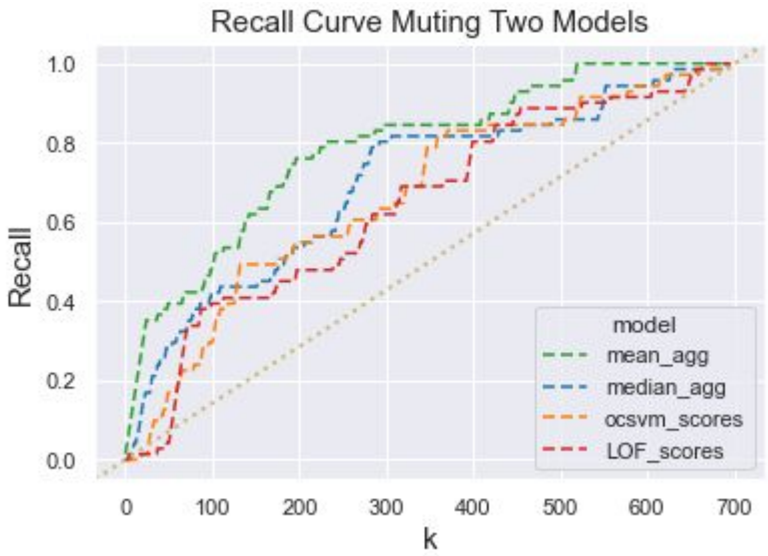


- xStream and LOF do better on this task
- iForest seems highly dependent on hyper-parameter tuning
- Difficult to assess performance without task-specific labeled data

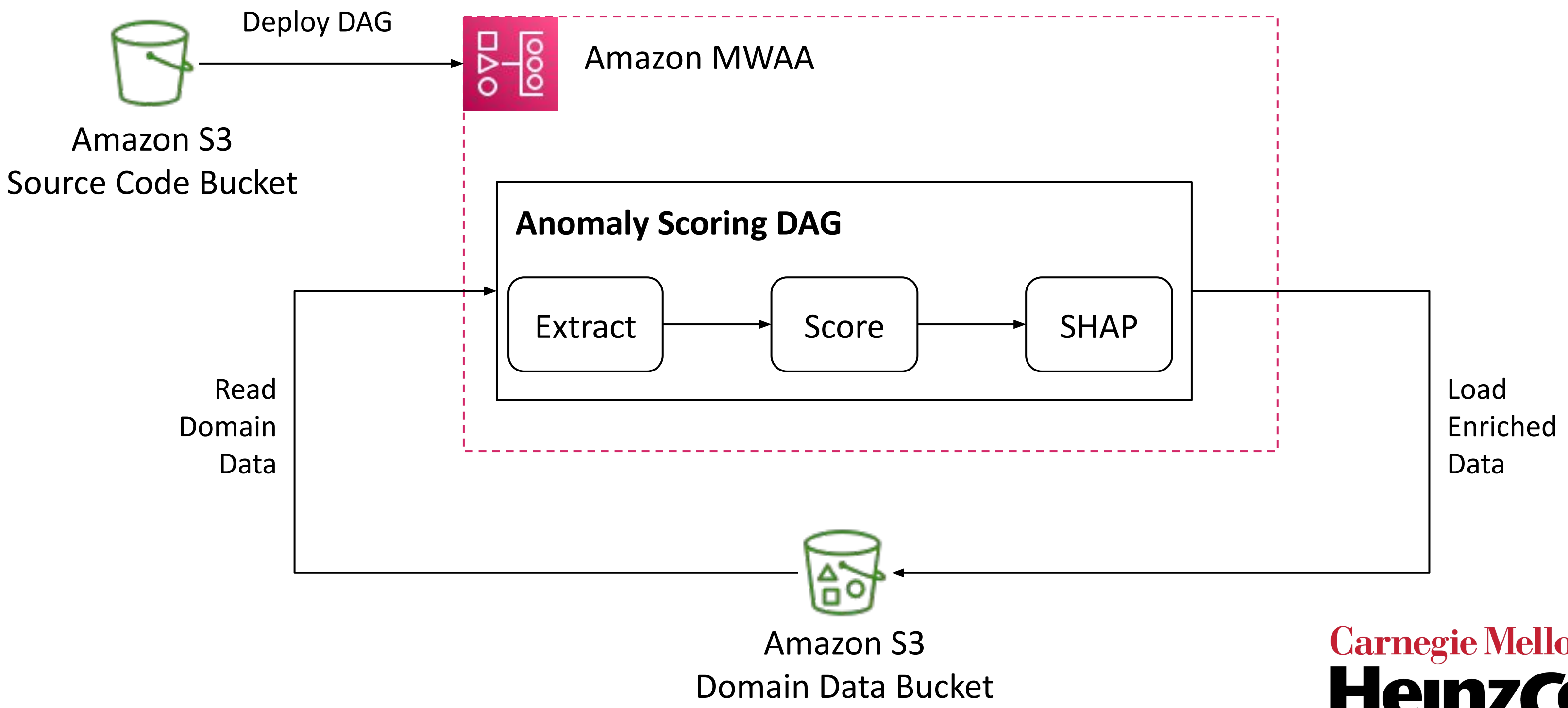
Isolation Forest xStream Local Outlier Factor

Rank Aggregation Resolves Sensitivity



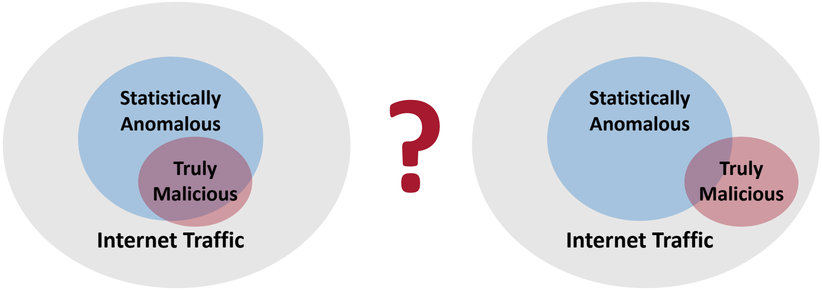


AWS Cloud Pipeline

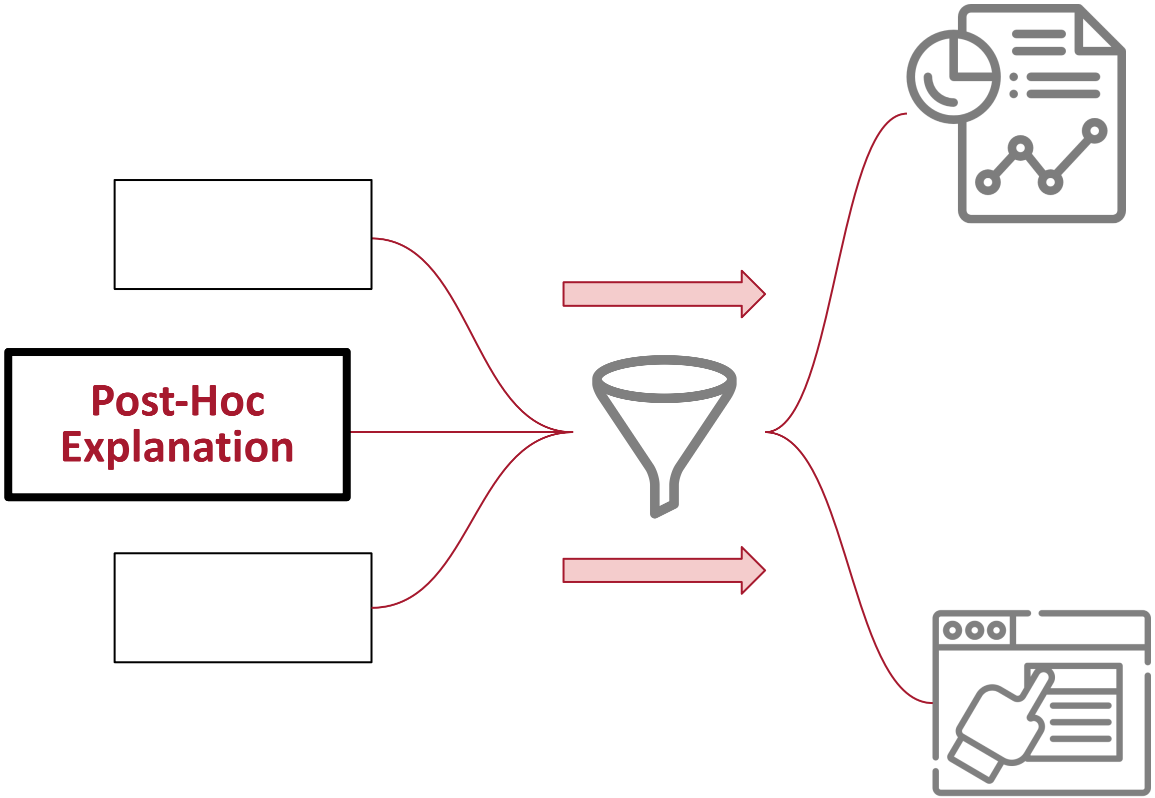


Malicious URL Conclusion:

Combining the anomaly scores of multiple models may improve performance when looking for malicious (spam) websites. However, we lack sufficient labeled data to determine the ability to detect malicious content targeting the Army network.

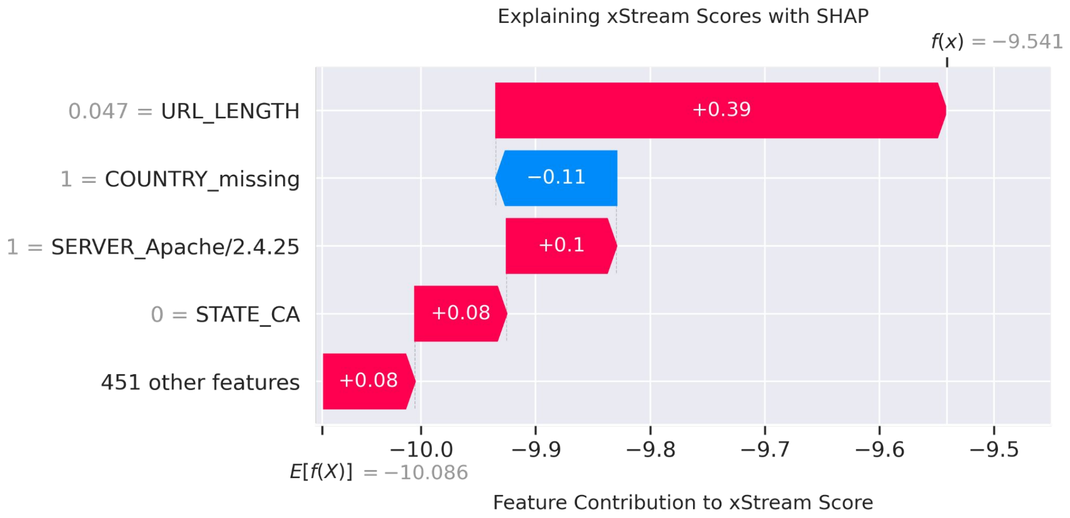


Carnegie Mellon University Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT


Post-Hoc
Explanation

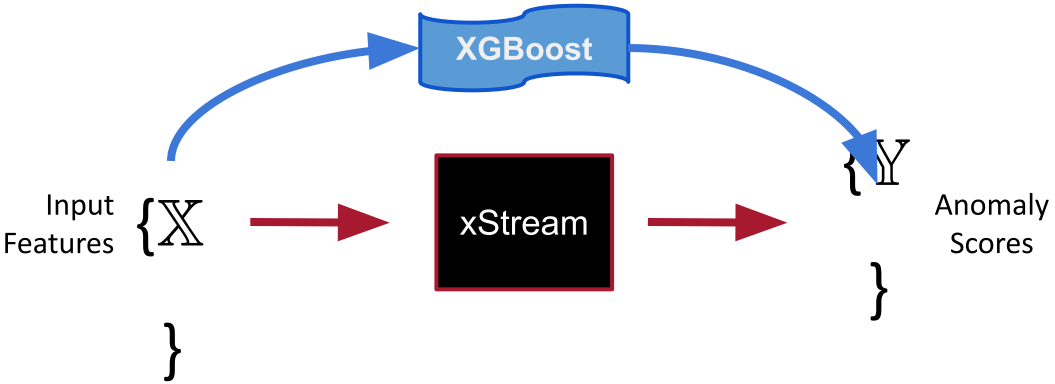
Post-hoc Explanation: Goals

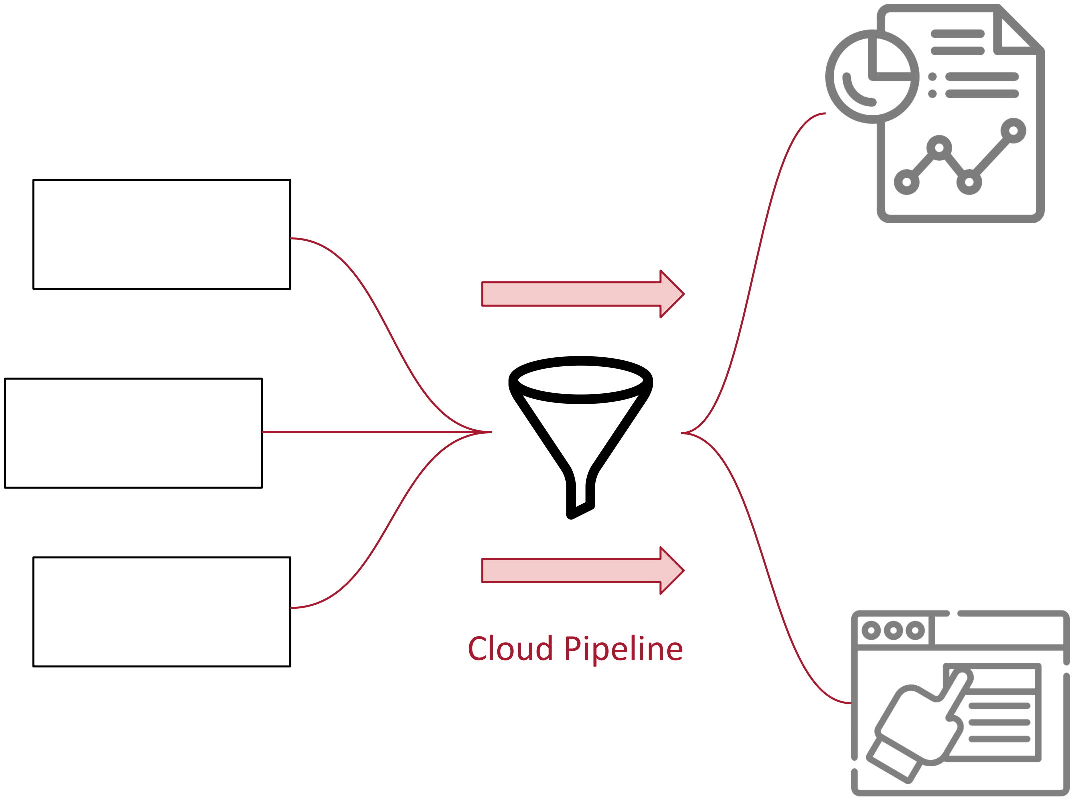
- Provide a human-interpretable explanation to anomaly scores
  - What factors contributed to the score?
  - What is different between this anomaly and non-anomalies?
  - What is different between this anomaly and *other* anomalies?

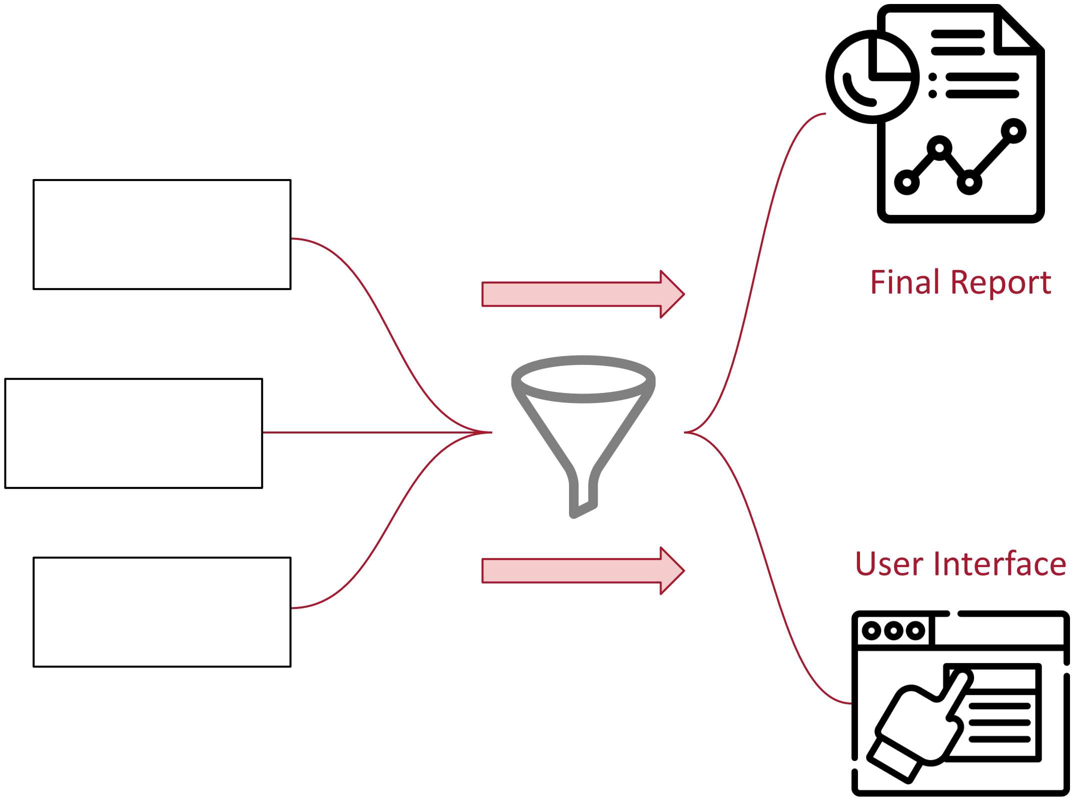


Post-hoc Explanation: Process

1. Collect unlabeled data
2. Generate xStream anomaly scores
3. Fit an XGBoost regression model to the data
4. Run SHAP on the regression model




Cloud Pipeline



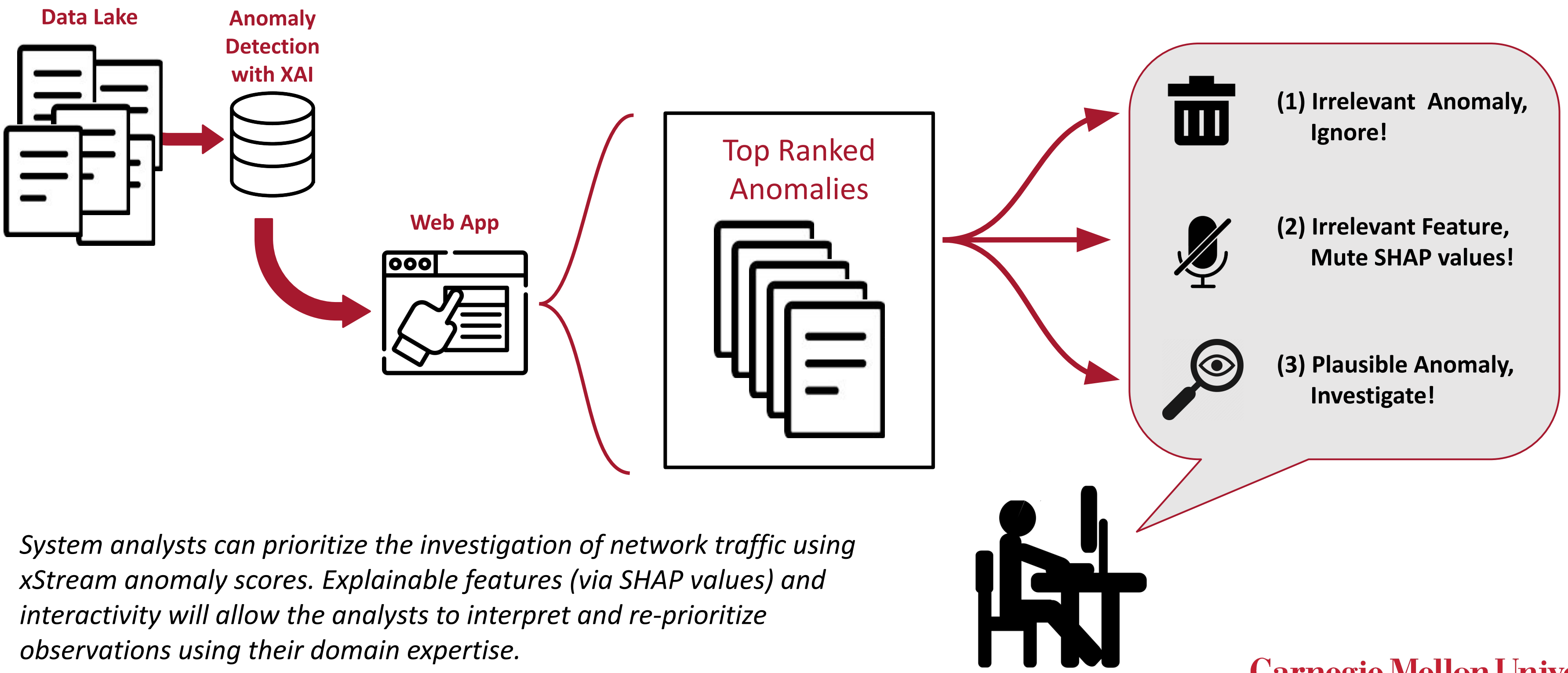
Final Report
User Interface

User Interface for Explainable Anomaly Detection

- **Goal**: Prototype an interactive interface demonstrating the combination of XAI and anomaly detection algorithms
  - Interoperable with NETCOM post-hoc analysis workflow and tools available to DSD
  - Ease of use for non-expert
- **Tool Selected**: R-Shiny (Web App)
  - + Rapid Development
  - + Complex Reactivity
  - + Deployable to Army Platforms (COEUS)
  - − Difficulties with Scaling

Carnegie Mellon University
HeinzCollege
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Proposed Workflow



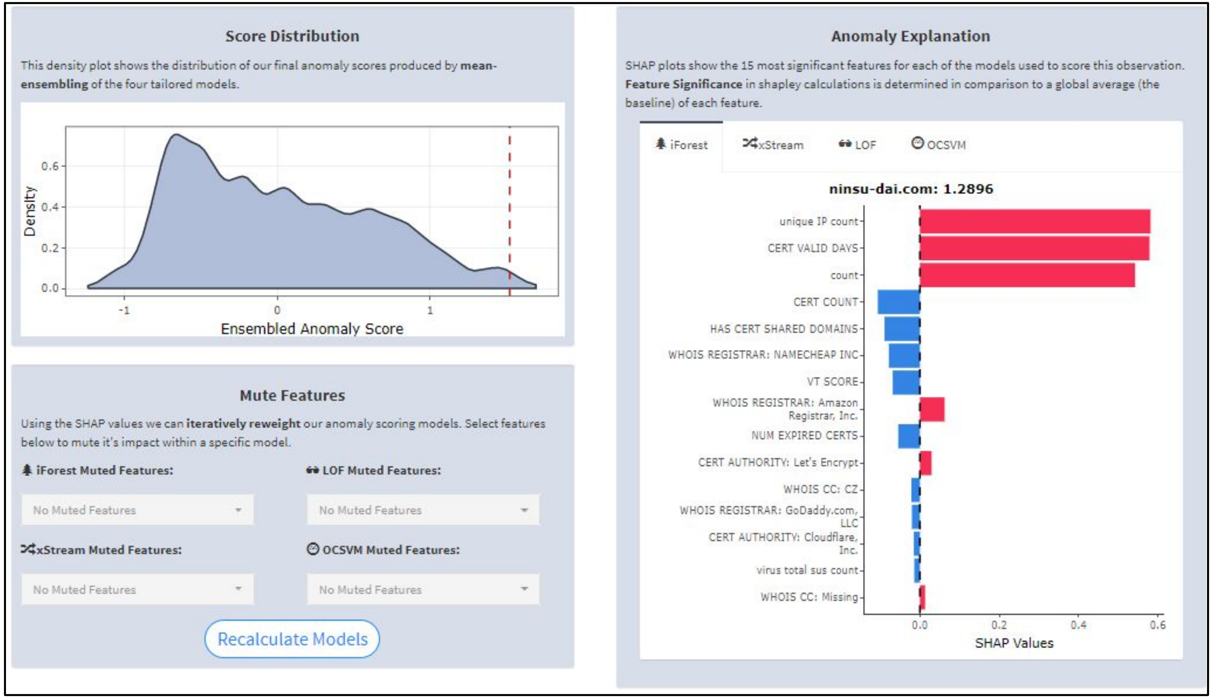
System analysts can prioritize the investigation of network traffic using xStream anomaly scores. Explainable features (via SHAP values) and interactivity will allow the analysts to interpret and re-prioritize observations using their domain expertise.

UI/X: Deployed Web App

https://jtmccorm.shinyapps.io/NETCOM_AD_Prototype/

UI/X Prototype: Tools for End-User

1. Rank Anomalies by Ensemble Scoring
2. Explain individual model outputs using SHAP values
3. Identify observations to investigate (or ignore)
4. Explore relationships between variables and scoring algorithms
5. Dynamically mute features preventing skew from irrelevant features
6. Download the dataset with annotations and updated scores

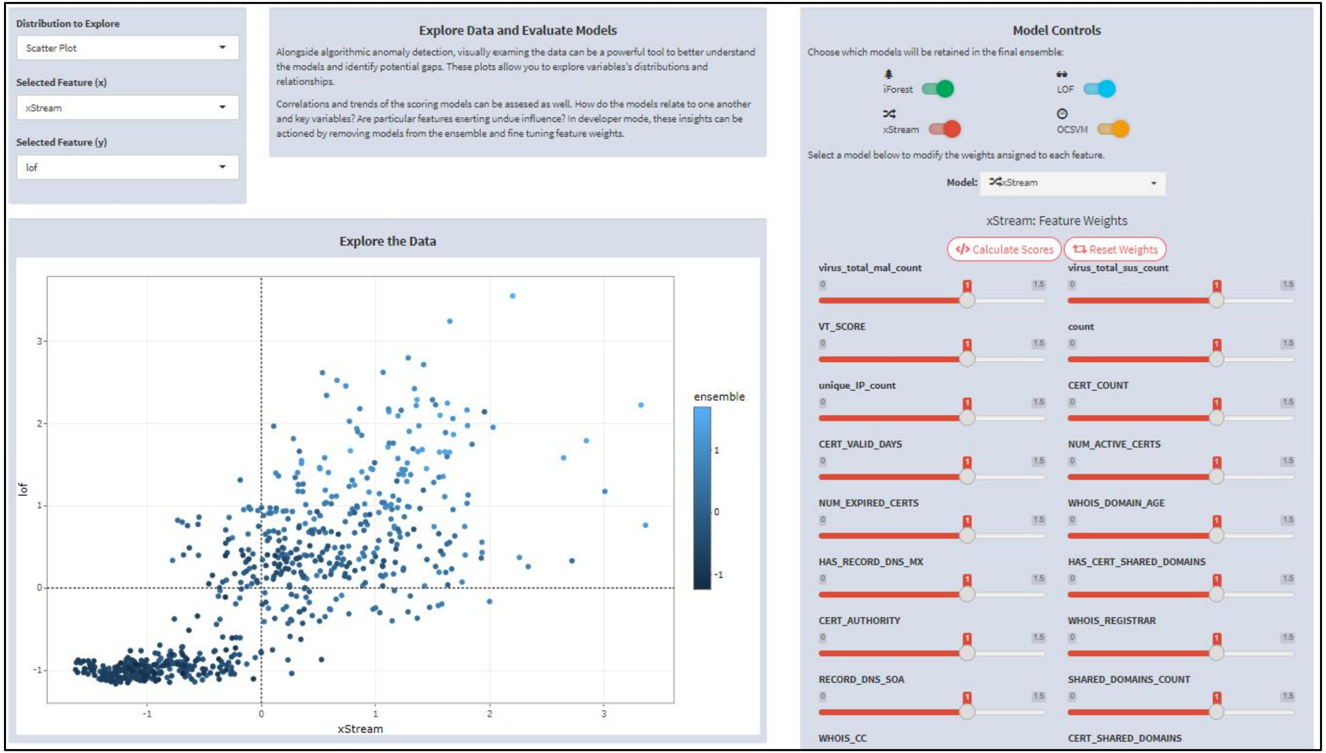


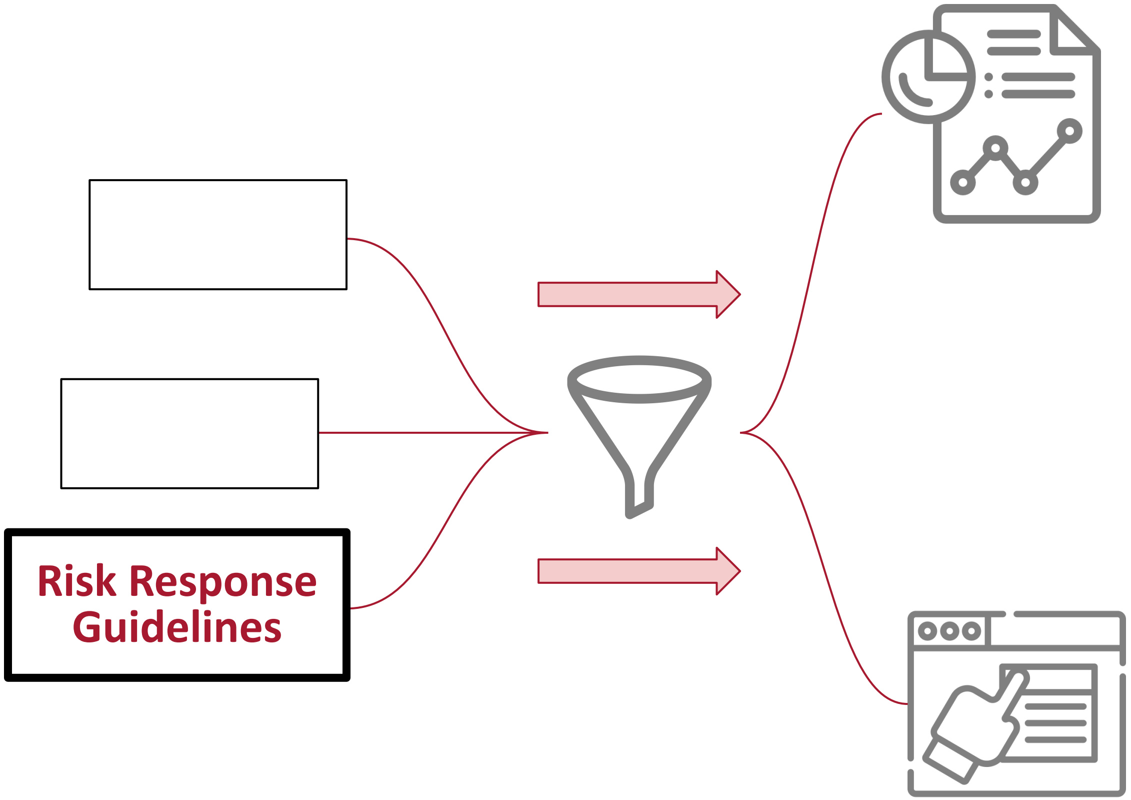
UI/X Prototype: Developer Access

1. Control which anomaly scoring methods are included within the ensemble.
2. Fine-tune feature weights of individual models (more granular controls than on user side)

User: jtmccorm

Password: 2017




Risk Response
Guidelines

Background & Goals

Background:

- Managing cyber risk is essential for NETCOM to achieve its strategic objectives
- Important to understand how NETCOM DSC-PIT fits into this process

Question: What role does our solution play in managing cyber risk?

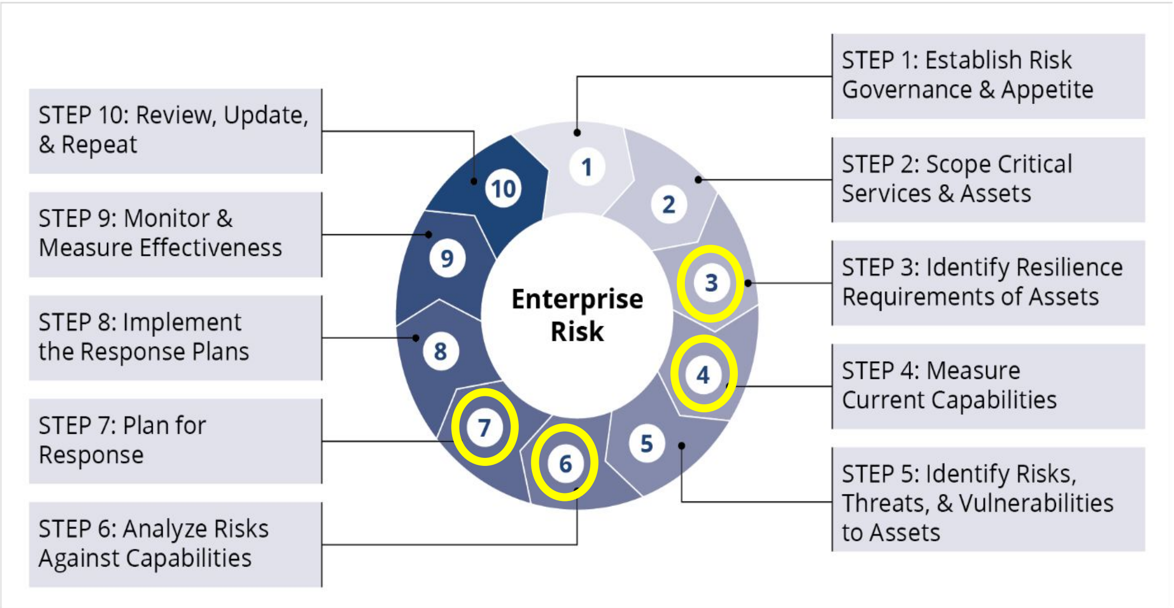
Goals:

- Show NETCOM DSC-PIT how to **translate anomaly scores into risk scores**
- Explain how this risk information feeds into cyber risk management practices that permeate through the Command → OCTAVE FORTE

Why OCTAVE FORTE?

Most critical steps for NETCOM DSC-PIT

- Provides an **enterprise-wide approach** to risk management
- Influenced by **standards that are used by the Army and DoD** for cyber risk management (ex. NIST RMF)
- Report touches on all steps, but some are more important than others for NETCOM DSC-PIT



https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=644636

Forming a Risk Score

(Ensemble) Anomaly Scores

- Indicate that there *may be* a malicious threat actor
- Indicate that there *may be* a vulnerability that is being exploited
- **Help to prioritize investigations**

Risk Components

- **Threat**
- **Vulnerability**
- **Asset/Exposure**
    - Confidentiality (C)
    - Integrity (I)
    - Availability (A)

"CIA Triad"

Impact Rating

Likelihood Rating

Inherent Risk Score

DELIBERATE RISK ASSESSMENT WORKSHEET

1. MISSION/TASK DESCRIPTION	**2. DATE** *(DD/MM/YYYY)*

3. PREPARED BY

a. Name *(Last, First, Middle Initial)*	b. Rank/Grade	c. Duty Title/Position
d. Unit	e. Work Email	f. Telephone *(DSN/Commercial (Include Area Code))*
g. UIC/CIN *(as required)*	h. Training Support/Lesson Plan or OPORD *(as required)*	i. Signature of Preparer

Five steps of Risk Management: (1) Identify the hazards (2) Assess the hazards (3) Develop controls & make decisions
(4) Implement controls (5) Supervise and evaluate *(Step numbers not equal to numbered items on form)*

4. SUBTASK/SUBSTEP OF MISSION/TASK	5. HAZARD	6. INITIAL RISK LEVEL	7. CONTROL	8. HOW TO IMPLEMENT/ WHO WILL IMPLEMENT	9. RESIDUAL RISK LEVEL
				How: Who:	
				How: Who:	
				How: Who:	
				How: Who:	
				How: Who:	

Additional entries for items 5 through 9 are provided on page 2.

10. OVERALL RESIDUAL RISK LEVEL *(All controls implemented)*:

☐ EXTREMELY HIGH ☐ HIGH ☐ MEDIUM ☒ LOW

11. OVERALL SUPERVISION PLAN AND RECOMMENDED COURSE OF ACTION

Risk Scoring Example: Data Exfiltration of Sensitive Payment Info

(Ensemble) Anomaly Score

Threat: Advanced Nation State Actor (NSA)

Vulnerability: Untrained soldiers,
weakness in blacklisting capabilities

Exposure: Payment information of multiple soldiers

Confidentiality: High
Integrity: Moderate
Availability: Low

Impact: Moderate

Likelihood: High

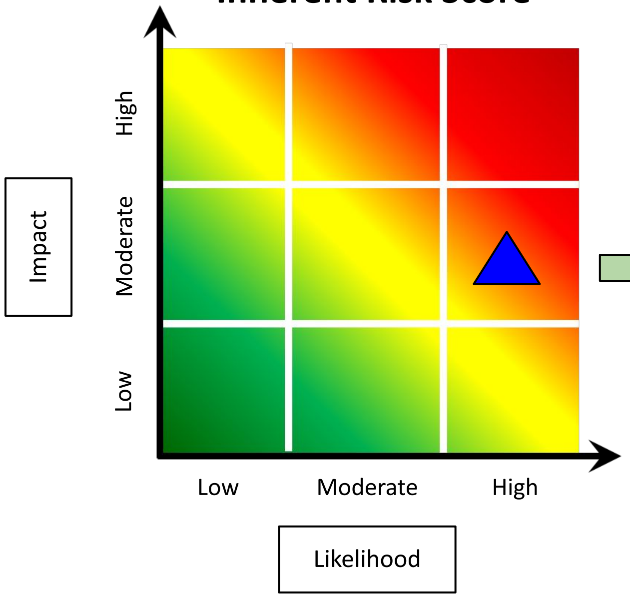
Inherent Risk Score
Moderate-High

Risk Scoring Example

Inherent Risk Score   Residual Risk Score

(prior to any control application)

Inherent Risk Score



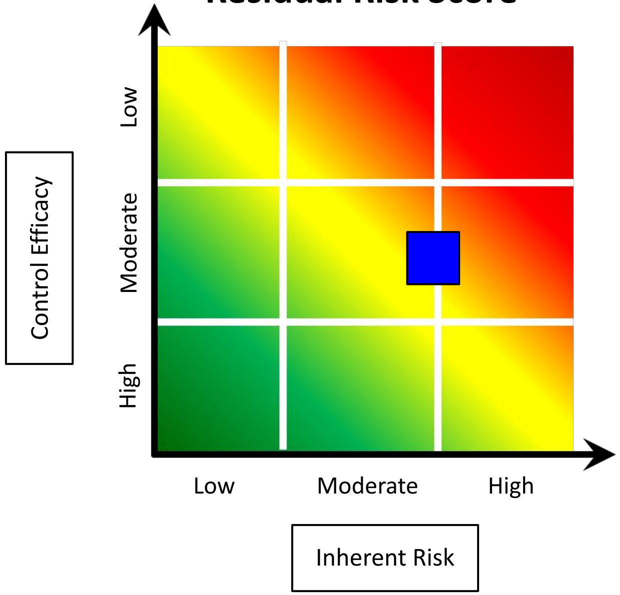
Response: Mitigate

Control:
Black-listing in IDS/IPS

Efficacy at Addressing Risk:
Moderate

(with controls considered, risk that remains)

Residual Risk Score



Residual Risk Score:
Moderate

Carnegie Mellon University
HeinzCollege
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Risk Response

- **Front-line Action:** Translating anomaly scores to risks scores gives NETCOM DSC-PIT a prioritized set of risks to address from the front line

- **Communication and Governance:** Our proposed governance structure shows NETCOM DSC-PIT how to effectively communicate relevant information up the cyber risk chain of command

- **Risk Appetite:** Forming a risk appetite statement allows NETCOM risk leaders to make informed, risk-based decisions around impact areas that affect strategic objectives

Sample Risk Appetite Statement

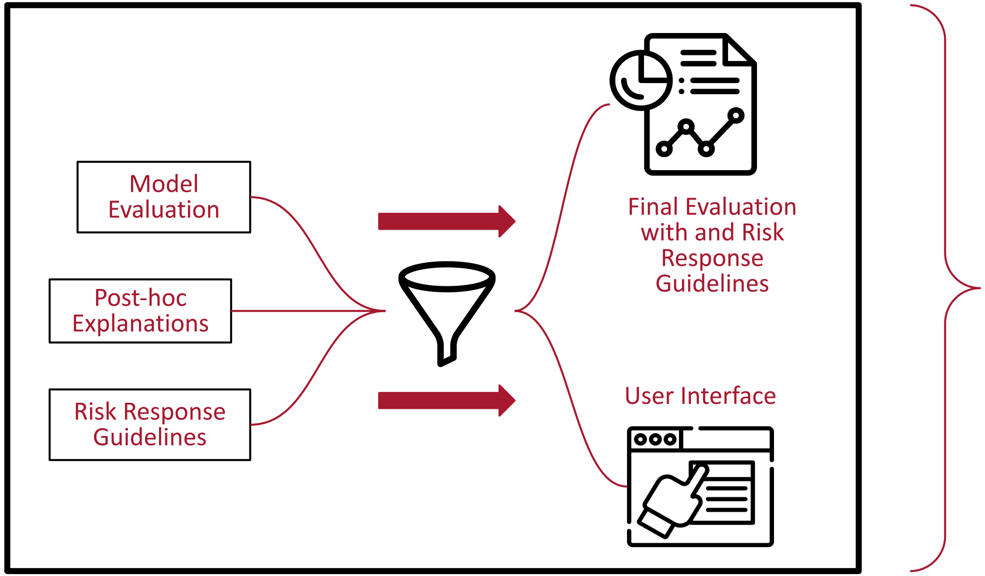
Escalation up Governance Tiers

	Data Exfiltration
Executive Attention (NETCOM Risk Leader)	• 20+ sensitive data records • 5+ classified records • 1+ top secret records
Management Attention (NETCOM DSD Risk Manager)	• 10+ sensitive data records • 1+ classified records
Front Line Attention (NETCOM DSC-PIT Risk Owner)	• 1+ sensitive data records

Example: **14** sensitive data records exposed

Carnegie Mellon University
HeinzCollege
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Our solution equips NETCOM with:



- **An understanding of how well xStream performs compared to other anomaly detection algorithms in relevant use cases**
- **An interactive user interface for prioritizing anomalies to investigate**
- **Risk response guidelines enabling effective action and informed decision making**

Future Work

- Investigate the impact of the temporal dimension on model performance (weekly, monthly, etc)
- Identify additional families of algorithms and approaches commonly used in this domain for comparison to this work
- Identify the specific "threat" and tune algorithms for specified task

Special Thanks to…

- Dan Costa
- **DSC-PIT**
  - MAJ Kevin Goulding
  - LTC Josiah Pickett
- Dr. Leman Akoglu
- Heinz College

Thank you!

References

Kreidler, N. (2019, December 12). *Project Sentinel - The Army Announces Cybersecurity Risk Management Framework Reform*. www.army.mil. Retrieved from https://www.army.mil/article/230900/project_sentinel_the_army_announces_cybersecurity_risk_management_framework_reform

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing* systems, 30.

National Institute of Standards and Technology (NIST). (2020, October 12). *Risk Management Framework for Information Systems and Organizations*. NIST. Retrieved from https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-37r2.pdf

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). " Why should i trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1135-1144).