Cyber Suite

montimage

Uptake of Innovative Security-as-a-Service Solutions

Hands-On Network Threat Detection and Response Using ACAS

Instructor: Edgardo Montes de Oca

Montimage

Co-funded by the European Union

Co-Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Digital Europe Programme. Neither the European Union nor Digital Europe Programme can be held responsible for them.

Agenda

01 Course Introduction

02 Module 1: Network Threats and Anomaly Detection Fundamentals

03 Module 2: Getting Started with ACAS

04 Module 3: Deep Packet Inspection & Feature Extraction

05 Module 4: Anomaly Detection with Rule-Based and Deep Learning Approaches

06 Module 5: Advanced Explainable AI (XAI) Analysis

07 Final Hands-On Assessment (Capstone)

08 Course Wrap-Up and Evaluation

Uptake of Innovative Security-as-a-Service Solutions

Course Introduction

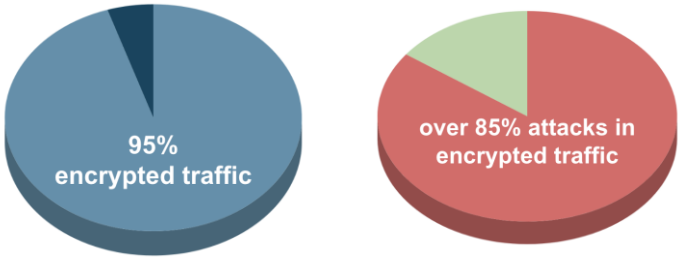
Co-funded by the European Union

Co-Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Digital Europe Programme. Neither the European Union nor Digital Europe Programme can be held responsible for them.

Context

- Network threats are evolving rapidly, and organizations need intelligent systems to detect anomalies before they become incidents.
- Encryption is rapidly becoming the norm
- Signature-based and rule-driven systems struggle to inspect encrypted payloads
- Decryption for inspection is resource-intensive, often impractical, and can raise privacy concerns

As of 2022



Source: State of Encrypted Attacks 2022 by Zscaler

Need for **AI-driven solutions**: Analyze metadata, traffic patterns, and behavioral signals without decrypting the content; Enable real-time, scalable, and privacy-preserving detection of threats

What you will learn

- Understand network threats and anomaly detection approaches (rule-based, machine learning)
- Navigate the ACAS platform and work with PCAP files
- Perform deep packet inspection and analyze traffic statistics and protocol hierarchies
- Extract and interpret flow-based features used for machine learning models
- Evaluate and compare rule-based vs. deep learning detection methods
- Apply Explainable AI methods (SHAP, LIME) to interpret model predictions

Who this course is for

This course suits IT and cybersecurity students, SOC analysts, network security engineers, and professionals seeking practical skills in AI-driven threat detection.

A basic understanding of networking (IP addresses, ports, protocols) and familiarity with cybersecurity concepts are recommended. No prior experience with machine learning or AI is required, every concept is introduced before it is used.

How the course is structured

The course is organised as five modules plus a final hands-on assessment. Each module is a short lesson. Modules 2-5 and Capstone are followed by an exercise or quiz.

Module	Example	Estimated Time
1	Network Threats and Anomaly Detection Fundamentals	1 hour
2	Getting Started with ACAS	2 hours
3	Deep Packet Inspection & Feature Extraction	3 hours
4	Anomaly Detection with Rule-Based and Deep Learning Approaches	4 hours
5	Advanced Explainable AI (XAI) Analysis	2 hours
Capstone	Final Hands-On Assessment	3 hours
Wrap-up	Course Wrap-Up and Evaluation	15 minutes

You can follow the course at your own pace; the full effort is less than **15 hours**, ideally spread over **several days**.

Recommended procedure to be followed

1. View each video of each module (or, if you prefer, go through the PowerPoint slide set and read the Word text).

2. Pause the video if some term or concept is not clear for you and search the web or ask a chatbot for more details. Active participation is important for acquiring higher knowledge.

3. For Modules 2-5 and Capstone, follow the instructions to do the exercises and quiz.

4. At the end, provide your feedback. This is important for us to be able to improve the course.

Hands-on exercises

We focus on several **web-based exercises** to explore and become familiar with key functionalities, including deep packet inspection (DPI), feature extraction, model evaluation, rule-based detection, threat prediction, and XAI-driven incident analysis.

Goals: to assess detection accuracy, usability, and performance under realistic network scenarios, gather participant feedback, and identify areas for improvement.

Link: https://strongcourage.notion.site/Hands-On-Network-Threat-Detection-and-Response-Using-ACAS-3881f0954df8802690bee935542f352b

Hands-On Network Threat Detection and Response Using ACAS

SCAN ME!

Assessment, Certification & Ethics

Assessment and certificate

Each module ends with a quiz; the pass mark is 7/10. After the capstone you complete a final quiz that draws on all the modules. Complete every module and the final assessment with the required grades and you earn a digital certificate of completion from the CyberSuite Academy.

A note on safety and ethics

When working with network traffic data, you may encounter sensitive information. Treat all data as confidential. The sample PCAP files provided for this course are sanitized datasets designed for educational purposes. Always follow your organization's data handling policies when working with real network captures.

Cyber Suite

montimage

Uptake of Innovative Security-as-a-Service Solutions

Module 1: Network Threats and Anomaly Detection Fundamentals

Co-funded by the European Union

Co-Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Digital Europe Programme. Neither the European Union nor Digital Europe Programme can be held responsible for them.

In this lesson

- Define network threats and explain the importance of threat detection for organizations
- Describe the different types of network attacks (DDoS, intrusion, malware, data exfiltration)
- Explain deep packet inspection (DPI) and its role in understanding network traffic
- Compare anomaly detection approaches: rule-based, statistical, and machine learning
- Understand the concept of Network Threat Detection and Response platforms

What are network threats?

A network threat is any malicious activity or event that could compromise the confidentiality, integrity, or availability of networked systems and data. Unlike vulnerabilities (which are weaknesses), threats are active attempts to exploit those weaknesses.

Organizations face an ever-growing landscape of network threats:

- Volume — millions of packets traverse networks every second
- Sophistication — attackers use advanced techniques to evade detection
- Speed — attacks can compromise systems in minutes or seconds
- Encryption — over 80% of traffic is now encrypted, hiding malicious content

Human analysts cannot manually inspect every packet; they need tools that surface the threats that matter. This is why **automated**, **intelligent detection systems** are essential.

Type of network attacks

montimage

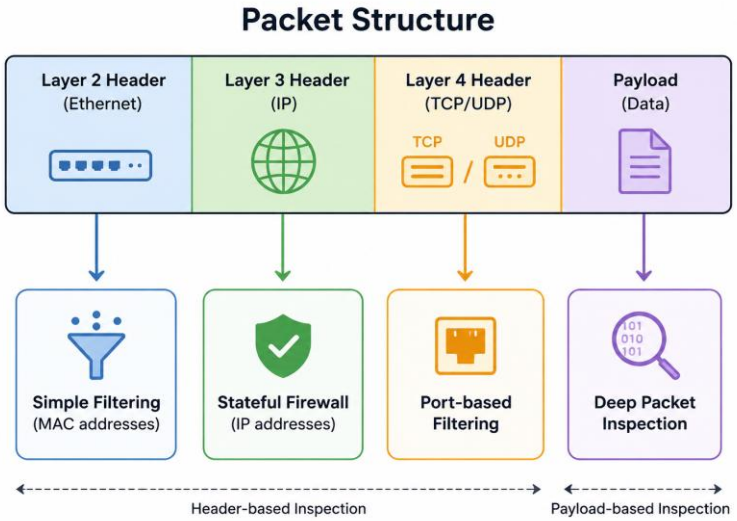
Cyber Suite

Attack Type	Description	Example	Network Indicator
DDoS	Overwhelm resources with traffic	SYN flood, UDP flood	High packet rates, unusual source distribution
Intrusion	Unauthorized access attempts	Brute force, exploitation	Failed logins, unusual port access
Malware	Malicious software communication	C2 beaconing, ransomware	Failed logins, unusual port access
Data Exfiltration	Unauthorized data transfer	DNS tunneling, covert channels	Large outbound transfers, unusual protocols
Reconnaissance	Network scanning and mapping	Port scans, service enumeration	Sequential port access, ICMP probes
Man-in-the-Middle	Traffic interception	ARP spoofing, SSL stripping	Duplicate MACs, certificate anomalies

Deep Packet Inspection (DPI)

Deep Packet Inspection is the examination of both the header and payload of network packets as they pass through an inspection point. DPI can:

- Identify protocols — recognize applications regardless of port (e.g., HTTP on port 8080)
- Extract content — analyze payload data for signatures or patterns
- Reconstruct sessions — follow conversations across multiple packets
- Detect anomalies — find unusual patterns that deviate from normal behavior



DPI examines all layers, including the payload, to make intelligent decisions about traffic classification and threat detection. DPI outputs in ACAS: **Traffic statistics**, **Protocol hierarchies**, **Bidirectional flow visualization** & **Session reconstruction**.

Anomaly detection approaches

Anomaly detection identifies patterns that deviate from expected behavior.

Two main approaches exist:

- **Rule-based (Signature) Detection**: matches traffic against known attack signatures, predefined patterns that indicate malicious activity.
- **Machine Learning Detection**: learns complex patterns from data, including deep learning approaches like CNNs and Autoencoders.

Approach	Known Attacks	Unknown Attacks	False Positives	Interpretability	Maintenance
Rule-based	Excellent	Poor	Low	High	High
Machine Learning	Excellent	Good	Variable	Low	Medium

Rule-based vs. ML-based detection comparison

1. Rule-based (Signature) Detection

Rule-based detection matches traffic against known attack signatures, predefined patterns that indicate malicious activity.

Advantages:

- Low false positive rate for known attacks
- Fast and computationally efficient
- Easy to understand and audit

Disadvantages:

- Cannot detect unknown (zero-day) attacks
- Requires constant signature updates
- Attackers can evade by modifying patterns

Example: *"Alert if more than 100 failed SSH login attempts from the same IP within 60 seconds".*

2. Machine Learning Detection

Machine learning models learn complex patterns from data, including deep learning approaches like CNNs and Autoencoders.

Advantages:

- Detects subtle, complex attack patterns
- Handles high-dimensional feature spaces
- Can identify unknown attacks

Disadvantages:

- Requires quality training data
- "Black box" — harder to interpret
- Computationally intensive

Example: *"A trained autoencoder flags traffic that cannot be reconstructed well, indicating anomalous patterns".*

Network Threat Detection and Response

Network Threat Detection and Response (NDR) is a category of security solutions that monitor network traffic to detect threats and enable rapid response.

Capability	Traditional IDS	NDR Platform
Detection method	Primarily signatures	Signatures + ML + behavior
Encrypted traffic	Limited	Advanced analysis
Response capabilities	Alert only	Alert + investigation + response
Threat hunting	Manual	Assisted/automated
Explainability	Rule matched	XAI insights

ACAS is a network threat detection platform that combines rule-based detection, deep learning models, and Explainable AI to provide comprehensive network threat detection and analysis.

Summary

- Network threats are malicious activities targeting the confidentiality, integrity, or availability of systems; they range from DDoS attacks to sophisticated intrusions and data exfiltration.
- Deep Packet Inspection (DPI) examines both headers and payloads to identify protocols, extract content, and detect anomalies; it is fundamental to understanding network behavior.
- Two anomaly detection approaches exist: rule-based (known signatures) and machine learning (learned patterns), each with distinct trade-offs.
- Modern NDR platforms like ACAS combine multiple detection approaches with AI/ML capabilities and Explainable AI to detect both known and unknown threats.

Uptake of Innovative Security-as-a-Service Solutions

Module 2: Getting Started with ACAS

Co-funded by the European Union

Co-Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Digital Europe Programme. Neither the European Union nor Digital Europe Programme can be held responsible for them.

In this lesson

- Describe what the ACAS platform is and the problem it solves
- Understand the platform's key features and capabilities
- Access the platform and create a user account
- Navigate the user interface and understand its main sections
- Upload and work with PCAP files for analysis

What ACAS is, and the problem it solves

montimage

Cyber Suite

The problem: Security analysts face overwhelming volumes of network data. Traditional tools generate alerts but provide little context about why traffic is suspicious or what to do about it. Analysts spend hours investigating false positives while real threats slip through.

The solution: ACAS combines multiple detection approaches with Explainable AI to:

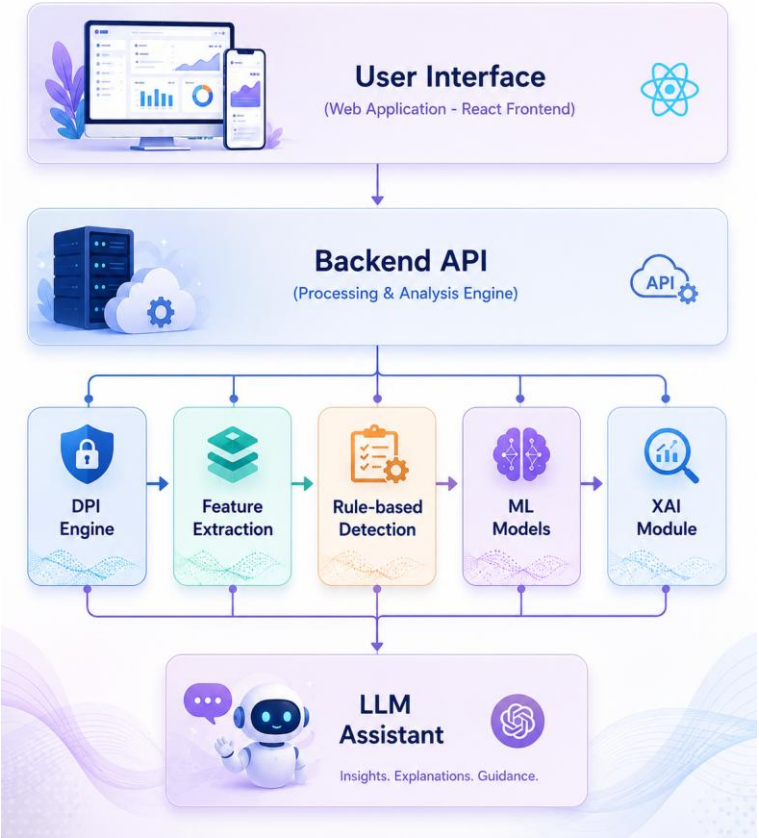
- Automatically analyze network traffic from PCAP files
- Extract meaningful features for machine learning
- Detect anomalies using both traditional and AI-based methods
- Explain why traffic is flagged as suspicious using SHAP and LIME
- Provide actionable recommendations through an LLM-powered assistant

ACAS, the Advanced Cybersecurity Analytics Service, is an AI-driven Network Threat Detection and Response platform developed by Montimage. It is designed for near real-time anomaly detection, threat prediction, and explainable analysis of network traffic, including encrypted traffic.

Platform architecture overview

Key components

- **DPI Engine**: performs deep packet inspection on uploaded PCAP files
- **Feature Extraction**: extracts 60+ flow-based features for ML analysis
- **Rule-based Detetion**: traditional signature matching
- **ML Models**: prebuilt detection models (CNNs, Autoencoders) for anomaly detection
- **XAI Module:** SHAP and LIME implementations for model interpretability
- **LLM Assistant**: provides natural language explanations and recommendations



Key features and capabilities

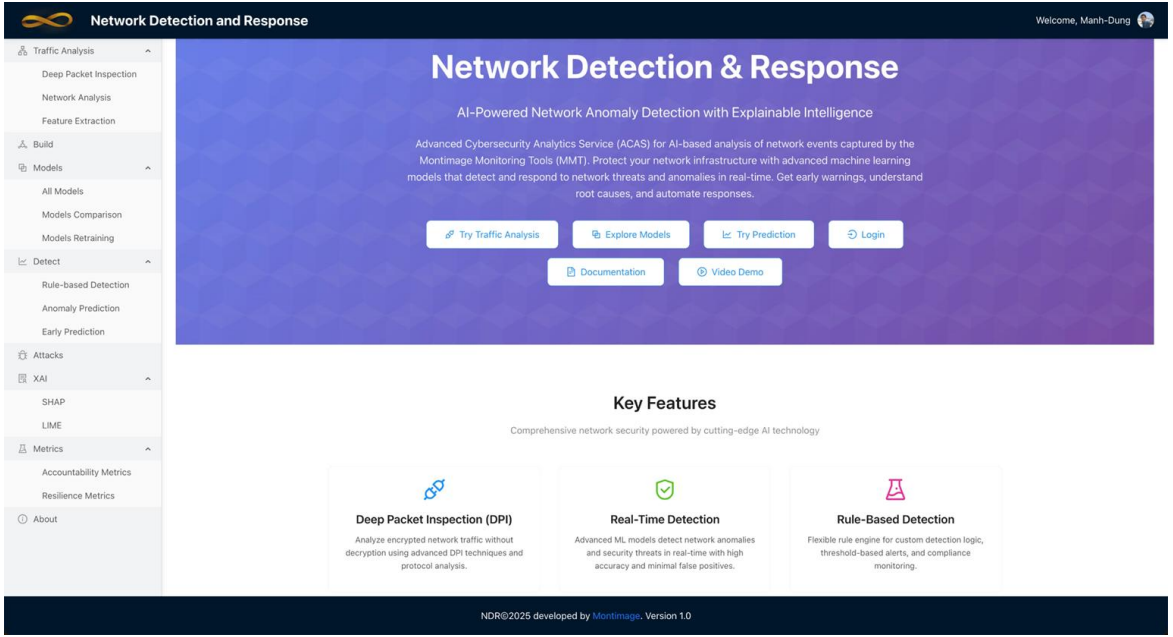
Feature	Description	Benefit
Deep Packet Inspection	Analyze packet contents, protocols, and flows	Understand network behavior in detail
Feature Extraction	Extract 60+ flow-based features automatically	Prepare data for ML without manual work
Rule-based Detection	Traditional signature matching	Catch known attack patterns reliably
ML (Prebuilt) Models	CNN and Autoencoder models for detection	Detect anomalies without training your own models
XAI Analysis	Global/Local feature importance & explanations	Understand which features drive predictions, Explain individual predictions
LLM Assistant	AI-powered recommendations	Get actionable mitigation guidance

Accessing the platform

The ACAS platform is accessible at: https://ndr.montimage.eu.

Creating an account:

1. Navigate to the platform URL in your web browser
2. Click "Sign in" using your Gmail/Github or "Sign Up"
3. Enter your email address and create a password
4. Verify your email if required
5. Log in with your credentials

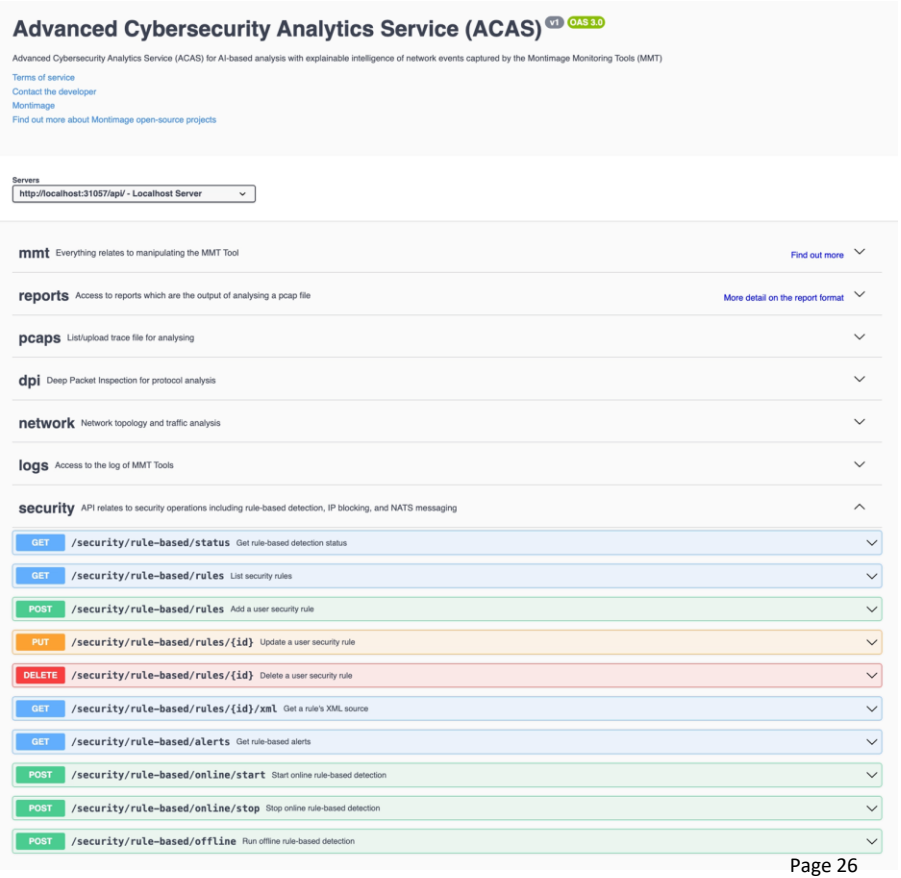


While registration provides access to all non-admin features, you can also explore some features without logging in. However, logging in is recommended to access the full functionality.

RESTful APIs

70+ Swagger-based APIs

- Manipulate the MMT tool
- Access reports and logs
- Upload trace file for analysis
- Manipulate models and the building model process
- Get the stats of the building model process
- Manipulate the prediction process and obtain results
- …



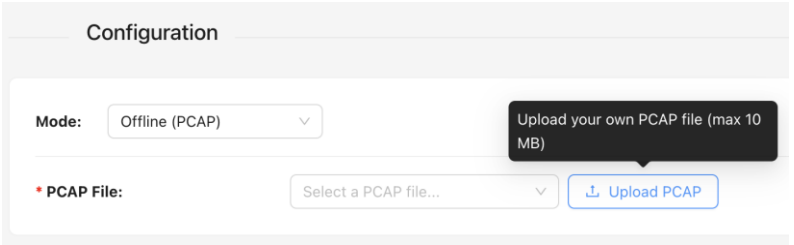
Working with PCAP files

PCAP (Packet Capture) files are the standard format for storing captured network traffic. ACAS analyzes these files (e.g., in .pcap or .pcapng) to perform all its functions.

File size limit: 10 MB maximum for uploaded files

The platform provides sample PCAP files for learning and testing, including:

- Normal traffic samples
- Various attack scenarios (DDoS, intrusion attempts, etc.)
- Mixed traffic for detection comparison



Best practices: Start with the provided sample files to learn the platform; Use smaller files (< 10 MB) for faster processing; Ensure PCAP files are properly formatted before upload

Hands-on exercise 1: Platform Exploration

1. Access the platform — Open https://ndr.montimage.eu in your browser
2. Register/Login — Create an account or log in
3. Explore the interface — Navigate through each section to understand the layout
4. Upload a sample file — Use one of the provided sample PCAP files
5. View DPI results — Examine the traffic statistics and protocol distribution
6. Note the features — Observe the extracted features for the uploaded file

Summary

- ACAS is an AI-driven NDR platform that combines DPI, feature extraction, multiple detection approaches, and Explainable AI to detect and explain network threats.
- The platform architecture includes a DPI engine, feature extraction module, ML models, XAI components, and an LLM assistant working together.
- Access the platform at https://ndr.montimage.eu and create an account to access full features.
- The interface is organized into sections for upload, DPI, feature extraction, detection, and explainability.
- PCAP files are the input format; the platform supports files up to 10 MB and provides sample files for learning.

Cyber Suite

montimage

Uptake of Innovative Security-as-a-Service Solutions

Module 3: Deep Packet Inspection & Feature Extraction

Co-funded by the European Union

Co-Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Digital Europe Programme. Neither the European Union nor Digital Europe Programme can be held responsible for them.

In this lesson

- Perform deep packet inspection on PCAP files and interpret the results
- Analyze traffic statistics including packet counts, byte volumes, and timing
- Understand protocol hierarchies and their distribution in network traffic
- Visualize bidirectional flows and identify communication patterns
- Extract flow-based features and understand their role in machine learning
- Interpret feature distributions and identify important features for threat detection

Deep packet inspection in ACAS

In Module 1, we introduced deep packet inspection (DPI) conceptually. Now we put it into practice using ACAS. When you upload a PCAP file, the DPI engine processes every packet to extract meaningful information about your network traffic.

What the DPI engine produces:

- Traffic Statistics: quantitative summaries of the captured traffic
- Protocol Hierarchies: breakdown of protocols present in the traffic
- Bidirectional Flows: visualization of communication between endpoints
- Temporal Patterns: how traffic volume changes over time

Traffic statistics

Traffic statistics provide a high-level view of the captured data. Several key metrics are listed in the table.

When analyzing statistics, look for:

- Unusual volumes — sudden spikes may indicate DDoS or data exfiltration
- Duration anomalies — very short or very long flows can be suspicious
- IP distribution — many sources to one destination suggests DDoS; one source to many destinations suggests scanning

Metric	Description	What it tells you
Total Packets	Number of packets in the capture	Overall traffic volume
Total Bytes	Sum of all packet sizes	Data transfer volume
Duration	Time span of the capture	Observation window
Packets/Second	Average packet rate	Traffic intensity
Bytes/Second	Average throughput	Bandwidth usage
Unique Source IPs	Distinct source addresses	Number of senders
Unique Dest IPs	Distinct destination addresses	Number of receivers
Unique Flows	Distinct 5-tuple combinations	Number of conversations

Protocol hierarchies

Protocol hierarchy analysis provides a high-level overview of network traffic by showing:

- Which protocols are present in the capture
- Their relative proportions
- Dominant communication patterns
- Potential anomalies or unexpected protocols

```
Frame (100%)
└── Ethernet (100%)
    └── IPv4 (98.5%)
        ├── TCP (75.2%)
        │   ├── HTTP (45.3%)
        │   ├── HTTPS/TLS (28.1%)
        │   └── SSH (1.8%)
        ├── UDP (22.8%)
        │   ├── DNS (15.2%)
        │   └── QUIC (7.6%)
        └── ICMP (0.5%)
    └── IPv6 (1.5%)
        └── ICMPv6 (1.5%)
```

Bidirectional flow visualization

A flow (or bidirectional flow) represents a complete conversation between two endpoints. ACAS identifies flows using the 5-tuple (source IP address, destination IP address, source port, destination port, protocol).

What to look for in flows:

- Asymmetric flows — large difference between sent/received may indicate exfiltration
- Long-duration flows — persistent connections could be C2 channels
- Unusual ports — unexpected services on non-standard ports
- Flow counts — many flows from one source could indicate scanning

Feature extraction for machine learning

ACAS automatically extracts over 60 flow-based features from PCAP files. As raw packets are not suitable for ML models, features transform packet data into numerical values that capture the essential characteristics of network behavior. Good features distinguish between normal and malicious traffic.

Feature Descriptions

Detailed metadata and explanations for each extracted feature

ID	Feature Name	Description	Type
1	`ip.session_id`	Session ID	numerical
2	`meta.direction`	Flow direction (0 means uplink and 1 means downlink)	categorical
3	`ip.pkts_per_flow`	Total number of IP packets	numerical
4	`duration`	Flow duration	numerical
5	`ip.header_len`	Total length of IP header	numerical
6	`ip.payload_len`	Total length of IP payload	numerical
7	`ip.avg_bytes_tot_len`	Average of total length of IP header	numerical
8	`time_between_pkts_sum`	Sum time between two flows	numerical
9	`time_between_pkts_avg`	Average time between two flows	numerical
10	`time_between_pkts_max`	Maximum time between two flows	numerical

Total 83 features < 1 2 3 4 5 ••• 9 > 10 / page

Hands-on exercise 2: DPI and Feature Extraction

montimage

Cyber Suite

1. Upload a PCAP file — Use a sample file from the platform
2. Examine traffic statistics — Note total packets, duration, and unique IPs
3. Review protocol hierarchy — Identify the dominant protocols
4. Explore bidirectional flows — Find the largest and longest flows
5. View extracted features — Examine at least 10 different features
6. Compare distributions — If multiple samples are available, compare feature distributions

Summary

- Deep packet inspection in ACAS produces traffic statistics, protocol hierarchies, bidirectional flow visualizations, and temporal patterns from PCAP files.
- Traffic statistics reveal volume, duration, and endpoint distribution; anomalies in these metrics can indicate threats.
- Protocol hierarchies show the breakdown of network protocols; unusual distributions warrant investigation.
- Bidirectional flows represent conversations between endpoints; asymmetric or long-duration flows may be suspicious.
- ACAS extracts 60+ flow-based features including packet lengths, inter-arrival times, flags, and rates for ML analysis.
- Feature importance varies by attack type; understanding which features matter helps interpret detection results.

Cyber Suite

montimage

Uptake of Innovative Security-as-a-Service Solutions

Module 4: Anomaly Detection with Rule-Based and Deep Learning Approaches

Co-funded by the European Union

Co-Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Digital Europe Programme. Neither the European Union nor Digital Europe Programme can be held responsible for them.

In this lesson

- Understand how rule-based detection works and configure detection rules
- Explain deep learning models for anomaly detection (CNNs and Autoencoders)
- Evaluate prebuilt models in ACAS and understand their configurations
- Interpret performance metrics: accuracy, precision, recall, F1-score
- Compare detection results from rule-based and deep learning approaches
- Analyze strengths and trade-offs of each detection method for different scenarios

Rule-based detection in ACAS

Rule-based detection uses predefined patterns (signatures) to identify known threats.

Rules work by continuously monitoring network traffic as a stream of events and checking whether specific patterns of those events match predefined conditions.

Each rule defines constraints on event types, attributes (like IPs, ports, or protocols), and sometimes timing or ordering between events.

When the observed traffic satisfies all the conditions of a rule, such as a sequence of actions happening in a specific order or multiple suspicious events occurring within a time window, the rule engine triggers an alert with a defined severity and message.

Examples of detection rules

montimage

Cyber Suite

Advantages:

- Predictable — you know exactly what triggers an alert
- Fast — simple pattern matching is computationally cheap
- Low false positives — for well-crafted rules targeting known attacks
- Auditable — rules can be reviewed and understood by humans

Limitations:

- No unknown attack detection — only catches what rules describe
- Maintenance burden — rules need constant updates for new threats
- Evasion — attackers can modify behavior to avoid matching rules

```
<beginning>
<!-- Property 15: Web scan with Nikto. User-Agent based
detection-->
<property value="THEN" delay_units="ms" delay_min="0"
delay_max="0" property_id="15" type_property="ATTACK"
     description="Nikto detection" >
   <event event_id="1" value="COMPUTE"
           description="Context: an user agent in the HTTP
header"
           boolean_expression="((http.method != '')
&amp;&amp; ((http.user_agent != '') &amp;&amp; (ip.src !=
ip.dst)))"/>
   <event event_id="2" value="COMPUTE"
           description="Trigger: Nikto detected. "
           boolean_expression="(#strstr(http.user_agent,
'Nikto') != 0)"/>
</property>
</beginning>
```

More rules here: https://github.com/Montimage/mmt-security/tree/main/rules

ML-based detection in ACAS

1. Convolutional Neural Networks (CNNs)

Originally designed for image processing, CNNs have been adapted for network traffic analysis.

How CNNs work:

Input Features → Convolution Layers (Pattern Extraction) → Pooling Layers (Feature Reduction) → Dense Layers (Decision Making) → Normal / Attack

CNN characteristics:

- Supervised learning (trained on labeled data)
- Binary or multi-class classification (normal vs. attack types)
- Good at detecting spatial patterns in features

2. Stacked Autoencoders (SAE)

Autoencoders are neural networks trained to reconstruct their input. They detect anomalies by identifying traffic that cannot be reconstructed well.

How Autoencoders work:

Input Features → Encoder → Latent Space → Decoder → Reconstructed Feature → Reconstruction Error → Low Error (Normal) / High Error (Anomaly)

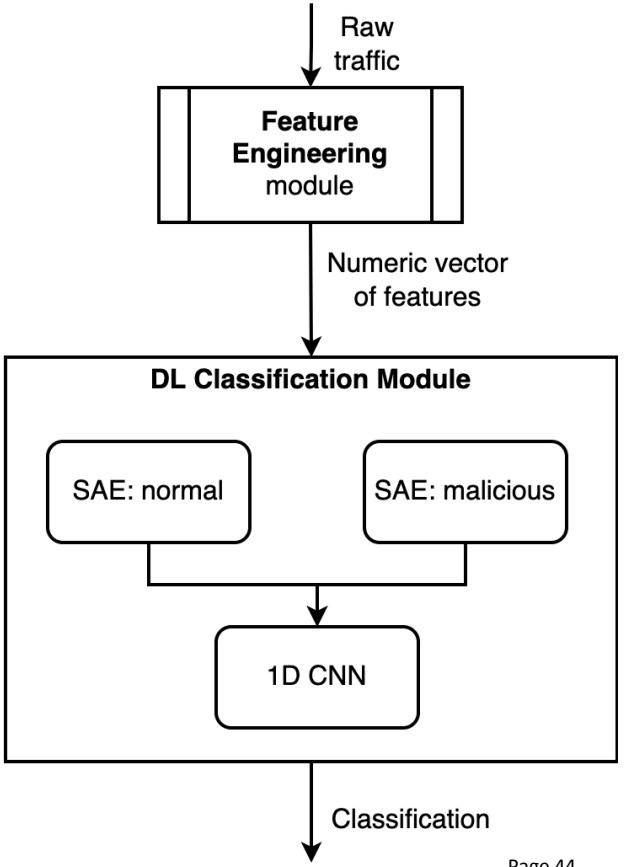
Autoencoder characteristics:

- Semi-supervised learning (trained mainly on normal data)
- Anomaly detection via reconstruction error
- Good at detecting novel attacks not seen in training

ML-based architecture

- Main functions:
  - **Traffic Feature Extraction**: 60+ flow-based features (protocol field values & statistical metrics) using MMT
  - **ML-Based Anomaly Detection**: combining CNN and SAEs
  - **Flexible Architecture**

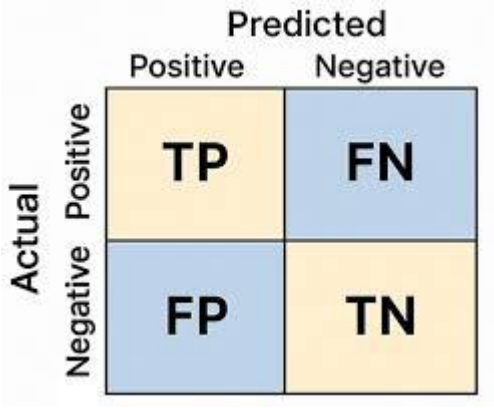
- Datasets: public (CSE-CIC-IDS2018: 99% in botnet detection) and private (captured from normal user activity and collected using our honeypot and attack generation tools)



Understanding performance metrics

To evaluate how well detection models perform, we use standard metrics, such as confusion matrix:

- TN = True Negative (correctly identified normal)
- FP = False Positive (normal flagged as attack)
- FN = False Negative (attack missed)
- TP = True Positive (correctly identified attack)



Interpreting metrics:

- High precision, low recall → Few false alarms, but missing attacks
- Low precision, high recall → Catching attacks, but many false alarms
- F1-Score → Good for imbalanced datasets; balances precision and recall

Metric	Formula	What it measures
Accuracy	(TP+TN)/(TP+TN+FP+FN)	Overall correctness
Precision	TP/(TP+FP)	How many alerts are real attacks
Recall	TP/(TP+FN)	How many attacks are detected
F1-Score	2×(Precision×Recall)/(Precision+Recall)	Balance of precision and recall

Strengths and trade-offs by scenario

Scenario	Best Approach	Reasoning
Known attack patterns	Rule-based	Fast, accurate, low false positives
Zero-day attacks	Autoencoder	Detects deviations from normal
Multi-class classification	CNN	Trained to distinguish attack types
High-value targets	All combined	Defense in depth
Resource-constrained	Rule-based	Lower computational requirements
SOC alert fatigue	Rule-based + XAI	Fewer alerts, explainable

Hands-on activity: Exploring anomaly detection techniques

montimage

Cyber Suite

1. Select a PCAP file
2. Run rule-based detection
3. Run ML-based detection
4. View metrics
5. Observe detected abnormal flows

Summary

- Rule-based detection matches traffic against predefined signatures; it is fast and precise for known attacks but cannot detect novel threats.
- CNNs are supervised deep learning models that classify traffic into categories based on learned patterns from labeled data.
- Autoencoders detect anomalies by measuring reconstruction error; traffic that doesn't match learned normal patterns produces high error.
- Performance metrics (accuracy, precision, recall, F1-score) help evaluate model effectiveness; the right balance depends on your use case.
- Comparing detection results reveals complementary strengths; rules catch known patterns, CNNs classify, and autoencoders find novelty.

Cyber Suite

montimage

Uptake of Innovative Security-as-a-Service Solutions

Module 5: Advanced Explainable AI (XAI) Analysis

Co-funded by the European Union

Co-Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Digital Europe Programme. Neither the European Union nor Digital Europe Programme can be held responsible for them.

In this lesson

- Understand why Explainable AI is essential for cybersecurity applications
- Apply SHAP (SHapley Additive exPlanations) to interpret model predictions globally and locally
- Apply LIME (Local Interpretable Model-agnostic Explanations) for individual prediction explanations
- Interpret feature importance visualizations and understand which features drive decisions

Why Explainable AI matters in cybersecurity

Machine learning models, especially deep learning, are often called “black boxes”, they make predictions, but we don't know why.

The **trust problem**:

- Analysts need to understand alerts to act on them
- "The model says attack" is not actionable
- False positives erode trust if unexplained
- Compliance and auditing require justification

The **accountability problem**:

- Who is responsible when a model makes a wrong decision?
- Can you defend your detection logic to stakeholders?
- Regulations increasingly require algorithmic transparency

Explainable AI (XAI) addresses these problems by **making model decisions interpretable**.
ACAS integrates XAI methods to help analysts **understand not just what was detected, but why**.

SHAP: SHapley Additive exPlanations

SHAP is based on game theory (Shapley values) and provides both global and local explanations. SHAP answers: "*How much did each feature contribute to this prediction?*"

How SHAP works

Think of features as "players" in a game, and the prediction as the "prize." SHAP calculates the fair contribution of each player to the final outcome. It provides global explanations, understand feature importance across all predictions.

SHAP output example

Global Feature Importance (SHAP values averaged across all samples):

- *flow_packets_per_second: 0.42*
- *total_fwd_packets: 0.31*
- *syn_flag_count: 0.28*
- *flow_duration: 0.19*
- *fwd_packet_length_mean: 0.12*

LIME: Local Interpretable Model-agnostic Explanations

LIME explains individual predictions by creating a simple, interpretable model around each prediction. LIME asks: "*If we slightly change the input features, how does the prediction change?*"

How LIME works

1. Take the instance to explain (e.g., one network flow)
2. Create perturbed samples by slightly modifying features
3. Get predictions for all perturbed samples
4. Fit a simple linear model to these results
5. The linear model's coefficients = feature importance

LIME output example

Explaining prediction for Flow #1247: Prediction ATTACK with probability 0.89.

Features supporting "ATTACK":

flow_packets_per_second > 5000 → *+0.32*
syn_flag_count > 1000 → *+0.24*
flow_duration < 10 seconds → *+0.18*

Features opposing "ATTACK":

total_bwd_packets > 100 → *-0.08*
fwd_packet_length_mean > 500 → *-0.05*

SHAP vs. LIME comparison

When to use each:

- Use SHAP when you need consistent, theoretically grounded explanations
- Use SHAP global to understand overall model behavior
- Use LIME for quick explanations of individual alerts
- Use both for high-stakes investigations

Aspect	SHAP	LIME
Theoretical basis	Game theory (Shapley values)	Local linear approximation
Scope	Global and local	Primarily local
Consistency	Mathematically consistent	May vary with perturbations
Computation	Can be slower	Generally faster
Interpretation	Additive feature contributions	Feature importance weights
Best for	Deep understanding	Quick local explanations

Hands-on activity: XAI analysis

1. Select a detected anomaly — Choose a flow flagged as an attack
2. Run SHAP analysis — Generate global explanations, which features are important
3. Run LIME analysis — Get local explanation for the flow
4. Formulate response — Based on insights, determine what action you would take

Summary

- Explainable AI addresses the "black box" problem, providing trust, accountability, and actionable insights for security analysts.
- SHAP uses game theory to calculate feature contributions; it provides both global (overall model) and local (individual prediction) explanations.
- LIME creates local linear approximations to explain individual predictions; it shows which features support or oppose the prediction.
- Feature importance visualizations (bar charts, force plots, dependence plots) help analysts quickly identify key factors in detections.
- Effective XAI usage follows a workflow: detect → explain → validate → act → document.

Cyber Suite

montimage

Uptake of Innovative Security-as-a-Service Solutions

Final Hands-On Assessment (Capstone)

Co-funded by the European Union

Co-Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Digital Europe Programme. Neither the European Union nor Digital Europe Programme can be held responsible for them.

Overview

This capstone assessment evaluates your ability to perform an end-to-end network threat analysis using the ACAS platform. You will apply all the skills learned throughout the course: deep packet inspection, feature extraction, anomaly detection, and explainable AI analysis.

Scenario

You are a SOC analyst at a mid-sized company. Your network monitoring system has captured traffic during a suspected security incident. Management needs a clear report explaining:

1. What threats (if any) are present in the traffic
2. How you detected and identified them
3. Why the traffic is considered malicious (with evidence)
4. What actions should be taken to respond

Your task is to analyze the provided PCAP file and produce a comprehensive analysis.

Suggested workflow

Step 0: Capture a small PCAP file containing attacks, such as brute-force attempts or DDoS traffic. Use:

- A freely available PCAP (https://github.com/Montimage/maip/tree/main/pcaps)

or

- capture your own PCAP file using wireshark (https://www.wireshark.org/).

Step 1: Initial analysis and DPI.

Step 2: Feature extraction analysis.

Step 3: Detection analysis using both rule-based and ML-based approaches.

Step 4: Explainability analysis using both LIME and SHAP methods.

Step 5: Final report on all findings and observations.

Summary

- Conducted an end-to-end analysis of network traffic using ACAS.
- Applied DPI, feature extraction, anomaly detection, and explainable AI techniques.
- Identified and investigated suspicious or malicious network activity.
- Validated findings using both rule-based and ML-based detection methods.
- Produced a comprehensive security report with evidence and response recommendations.
- Demonstrated practical SOC analyst skills for real-world incident investigation.

Cyber Suite

montimage

Uptake of Innovative Security-as-a-Service Solutions

Course Wrap-Up and Evaluation

Co-funded by the European Union

Co-Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Digital Europe Programme. Neither the European Union nor Digital Europe Programme can be held responsible for them.

Course wrap-up

Congratulations on completing **Hands-On Network Threat Detection and Response Using ACAS**!

What you learned from this course

- **Recognize and analyze network threats**, including DDoS attacks, intrusions, malware communications, and data exfiltration activities.
- **Perform Deep Packet Inspection (DPI)** to investigate traffic statistics, protocol hierarchies, and bidirectional network flows.
- **Extract and interpret flow-based features** used by AI models for network anomaly detection.
- **Compare rule-based and AI-driven detection approaches**, understanding their strengths, limitations, and practical applications.
- **Apply Explainable AI (XAI) techniques**, including SHAP and LIME, to interpret and justify machine learning predictions.
- **Conduct end-to-end threat investigations**, from packet capture analysis to evidence-based security reporting and incident response recommendations.

We Value Your Feedback!

Please take a few minutes to complete our short course-evaluation survey. Your honest feedback helps us improve the lessons, labs, and the ACAS tool itself for future learners.

What the survey asks about

- The clarity and usefulness of each module.
- Whether the hands-on labs were easy to follow and worked as expected.
- The pacing and overall difficulty of the course.
- Any technical issues you encountered with the platform.
- Suggestions and features you would like to see.

Take the survey

Please complete the evaluation form. The survey is anonymous and takes about five minutes. Thank you for helping us make the course better and good luck applying your new network threat detection and response skills, responsibly and only on systems you are authorised to test.

Thank you for your attention!

Montimage

Edgardo Montes de Oca, Manh Nguyen

manhdung.nguyen@montimage.eu
edgardo.montesdeoca@montimage.eu

https://www.montimage.eu/

Co-funded by the European Union

Co-Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Digital Europe Programme. Neither the European Union nor Digital Europe Programme can be held responsible for them.