# CIS 350: AI for Cybersecurity

# Week 6 Activity Worksheet: Intrusion Detection and Anomaly Analysis

## Activity Overview

In this activity you will study a guided lesson on intrusion detection and anomaly analysis, then answer questions to deepen your understanding of how security tools tell normal activity from an attack. You will connect the ideas to Lab 5/6, where you scaled features and tuned the contamination rate of an Isolation-Forest-style anomaly detector. Finally, you will design a practical detection tuning and alert-triage checklist for a small business.

### Warm-Up (5 min): Predict First

Before you study, make a quick prediction and discuss it with a partner. Do NOT look anything up yet - this is just your gut instinct.

A security detector flags 100 alerts per day, and only about 2 of them turn out to be real attacks. In one or two sentences: is that a good detector or a bad one? What do you think goes wrong when 98 of the 100 alerts are false alarms? Jot your prediction below, then compare with your partner. You will revisit this exact scenario in Task 3.

### Task 1: Study the Lesson

**Resource:** Hacker101 Lessons (hacker101.com/lessons) or NIST anomaly-detection guidance (nvlpubs.nist.gov)
Time: about 15 minutes

As you study, take notes on:
- What an intrusion detection system (IDS) does and who reads its alerts
- The difference between signature-based and anomaly-based detection
- Why anomaly detectors raise false alarms and what that costs a business
- Any example of a detector missing a brand-new (novel) attack

### Task 2: Guided Questions

Q1: In your own words, what does an intrusion detection system (IDS) do, and why can't a human just watch the logs instead?

Q2: Explain the difference between signature-based and anomaly-based detection. Give one strength of each.

Q3: Why does feature scale matter for an anomaly detector like Isolation Forest? What can go wrong if you skip scaling?

Q4: A detector flags 100 sessions per day, but only about 2% are real attacks. How many false alarms is that, and why is that a problem for the analysts?

Q5: Unsupervised anomaly detection can catch zero-day attacks that supervised models miss. Why does that NOT make it always the better choice?

## Task 3: Analyze the Lab 5/6 False-Alarm Numbers

In Lab 5/6 you scaled features before clustering and tuned the contamination rate of an Isolation-Forest-style anomaly detector. Recall this scenario:
- The detector flags about 100 sessions per day
- About 2% of flagged sessions are real attacks (roughly 2 per day)
- The other roughly 98 flags per day are false alarms
- Each flagged session takes an analyst time to investigate

**1.** If an analyst needs 10 minutes to investigate each flagged session, how many hours per day do the 100 alerts cost? How much of that time is spent on false alarms?

**2.** Suppose you tune contamination STRICTER so the detector flags only 20 sessions per day. What is the risk to the business, and which of the 2 real attacks might you now miss?

**3.** How would you MEASURE whether the detector is actually preventing attacks (not just generating alerts)? Name at least two things you would track over time.

## Task 4: Design a Detection Tuning and Alert-Triage Checklist

You are the security lead at a small online retailer. Design a checklist of 5-7 steps for setting up and tuning an anomaly detector and for triaging its alerts. Your checklist should address feature scaling, choosing the contamination rate, validating clusters, and handling false alarms.

**Your Checklist (use a numbered list):**

**Explain your choices:** Why did you order the steps this way? How does your checklist balance catching real attacks against overwhelming your analysts with false alarms?

## Task 5: Reflection

Q1: What surprised you most about how anomaly detectors work or fail? Why?

Q2: If you found a suspicious cluster in your data, how would you decide whether it is a real group of attackers or just a data artifact? Describe your check.

Q3: What is one thing you will remember about the trade-off between false alarms and missed attacks the next time you tune a security tool?

## Grading Rubric (20 points)

| Criterion | Description | Points |
|---|---|---|
| Completion | Lesson studied; all 5 tasks and questions attempted | 8 |
| Understanding | Q&A answers demonstrate learning about IDS and anomaly detection | 7 |
| Application | Detection/triage checklist is practical, well-reasoned, and explained | 5 |

**Total: 20 points**

Scoring: Completion (8) + Understanding (7) + Application (5) = 20 points

## How to Submit

1. Complete all tasks in this worksheet
2. Ensure all questions are answered thoughtfully
3. Submit this PDF on Canvas by [DUE DATE]

Questions? Post in the course forum or attend office hours!