HANDOUT 2 - WEEK 2 READING

Introduction to Data and AI

CIS 350: AI for Cybersecurity - Smart Defense for the Digital Business

Colorado State University - Computer Information Systems

How to use this reading

This handout supports Week 2. Read it before and after class. It explains, in plain language, what AI and machine learning really are, how rules differ from learning, where security data comes from, and how a dataset is built from rows, features, and labels. The review questions at the end map directly to Quiz 2 and Lab 2.

Estimated reading time: 20-25 minutes. No prior technical background needed.

1. Data: The Fuel of AI

Last week we asked what we protect. This week we look at the engine behind modern defense: data, and the artificial intelligence that learns from it. Every login, email, and network connection leaves a trail of data - and that data is exactly what AI uses to spot attacks.

The single most important idea in this handout is simple: AI is only as good as the data it learns from. Feed a model clean, accurate, well-organized data and it can be remarkably useful. Feed it messy or mislabeled data and it will be confidently wrong. Security professionals spend most of their AI effort collecting, cleaning, and organizing data long before any clever algorithm runs.

2. AI and Machine Learning, in Plain Language

Artificial intelligence (AI) means getting computers to do tasks that normally need human intelligence - recognizing faces, understanding language, detecting fraud, or recommending products. It is not magic and not conscious; it is math finding patterns in data. AI is the broad field; machine learning is the part driving today's boom.

What AI is NOT (cutting through the hype)

- NOT conscious or 'thinking' - it has no understanding or intent.
- NOT always right - it makes confident mistakes, especially on bad data.
- NOT magic - every result traces back to data and math you can inspect.
- NOT a replacement for judgment - in security, humans stay in the loop.

Machine learning: learning from examples

Traditional software follows rules a human writes by hand. Machine learning flips this: we show the computer many labeled examples and it discovers the pattern itself. Show a model 10,000 emails marked 'spam' or 'not spam' and it learns to tell them apart - nobody writes 'if the subject says X, mark it spam.' The model finds the clues on its own, and more good examples usually means a better model.

> **How a model 'learns' (no math)**
>
> Training is just a loop: GUESS (the model predicts), CHECK (compare the guess to the real answer, the label), ADJUST (nudge its internal numbers to be a little more right). Repeat over thousands of examples until the predictions are good. That is all 'training' really means - and it is why labeled examples are so valuable.

3. Rules vs. Learning From Data

There are two ways to make software decide. Rule-based systems use conditions a human writes ('if 5 failed logins, lock the account'). Machine learning systems learn patterns from data. Neither is automatically better - you fit the tool to the problem.

When each approach wins

- Rules win when the problem is simple, stable, and easy to state - and you want full transparency.
- Rules struggle against creative attackers: block 'FREE' and they write 'F-R-E-E' or 'Fr33', forcing endless new rules.
- Learning wins when patterns are complex, shifting, and hidden in lots of data - like phishing that changes daily.
- Machine learning needs good, labeled data and is harder to fully explain.
- In practice, real systems combine both: fast rules as a first filter, plus an adaptive model.

4. The Four Styles of Machine Learning

- Supervised - learn from labeled examples (spam vs. not spam). The focus of this course.
- Unsupervised - find patterns with no labels (group similar logins or traffic).
- Semi-supervised - learn from a few labels plus lots of unlabeled data.
- Reinforcement - learn by trial, reward, and penalty (like a game).

For the next several weeks we focus on supervised learning, because it matches the labeled spam example and the way most security detection works today.

5. Everyday AI in Business

You already rely on AI every day. Recommendations ('customers also bought...') learn from what huge numbers of people viewed and bought. Fraud alerts learn your normal spending and flag the unusual. Spam filters learn from labeled mail. Chatbots learn language patterns to answer questions. Each is the same idea: learn a pattern from data, then act on new cases.

Why this course pairs AI with security

The same techniques that recommend products or flag fraud also detect phishing, intrusions, and anomalies. And the flip side matters: attackers use AI too - to write convincing phishing at scale, generate malware variants, and create deepfakes. That dual-use reality is why understanding both AI and security is so valuable.

6. Where Security Data Comes From

Security data comes from five main sources, and real investigations usually combine several of them.

- System logs - timestamped records of who did what, when (logins, access, changes). The backbone of most investigations.
- Email - senders, subjects, links, and attachments; powers phishing and spam detection.
- Network traffic - records of connections that can reveal data leaving to unknown servers or unusual volumes.
- Endpoints - the devices themselves (laptops, phones, servers): processes, files, and settings.
- User activity - clicks, downloads, and login times that form a behavior 'fingerprint'.

Raw data is messy and not ready for AI as-is. Teams collect it, clean it (fix errors and missing values), label it (mark examples such as 'phishing' vs. 'legitimate'), and pick features - the clues a model will learn

from. Privacy matters throughout: this data is sensitive and must be handled responsibly.

7. The Anatomy of a Dataset

A dataset is just a smart spreadsheet: a table of rows and columns. Each ROW is one example or record (one email, one login, one transaction). Each COLUMN is one piece of information about every row.

Features vs. label - the key distinction

- Features are the input clues the model looks at (number of links, attachment yes/no, sender mismatch, urgent words).
- The label is the answer we want to predict (spam or not spam).
- Supervised learning means learning to predict the label from the features - 'features in, label out.'
- Features come in types: numeric (counts, amounts), categorical (yes/no, country), and text-derived (contains 'urgent').
- Choosing and creating good features is called feature engineering, and good features matter more than fancy algorithms.

Structured vs. unstructured data

Structured data already fits neatly in rows and columns (a login log, a transactions table). Unstructured data does not (the raw text of an email, an image, a PDF). Much security data is unstructured, so we extract features to give it structure - for example, turning a raw email into counts of links and flags for attachments. During training a model also sees the real labels so it can check and improve; later it predicts labels for brand-new, unlabeled data.

8. Data Quality: Garbage In, Garbage Out

A model trained on bad data learns the wrong lesson - confidently. In security that means missed attacks or endless false alarms, so checking data quality is essential.

Common data-quality problems

- Missing values - blank fields the model cannot read.
- Class imbalance - one class is rare (e.g., 1 attack per 1,000 events).
- Mislabeled data - a spam email marked 'ham' teaches the wrong answer.
- Bias - data that does not represent the real world (e.g., attacks only from one place).

Why imbalance makes accuracy lie

If only 1 in 1,000 events is an attack, a lazy model that always says 'normal' is 99.9% accurate - yet it catches ZERO attacks. High accuracy can completely hide failure on the cases that matter most. That is why we use measures like precision and recall (Week 4) instead of accuracy alone, and why we balance and relabel data.

To improve data quality: clean it, balance rare classes by collecting more examples or resampling, double-check important labels, document where the data came from, and re-check often because data drifts as attackers and users change over time.

9. Visualizing Data to Spot Patterns

Before building any model, analysts look at the data, because a picture surfaces problems faster than rows of numbers. Match the chart to the question:

- Bar chart - compare counts across categories (which attack types are most common).
- Line chart - track change over time (logins per hour; a spike can signal an attack).
- Scatter plot - reveal relationships and outliers (a few unusually large transactions).
- Histogram - show the spread of one value (how many links emails tend to have).

A sudden spike, a lonely outlier, or a tall bar often points straight to something worth investigating. In Lab 2 you will load a real phishing-email dataset, count phishing vs. legitimate mail, and chart the features that separate them.

Key Terms

- AI / machine learning - doing intelligence-like tasks / learning patterns from labeled examples.
- Rule-based vs. learning - hand-written conditions vs. patterns learned from data.
- Row / record - one example in a dataset (one email, one login).
- Feature - an input clue; Label - the answer to predict; feature engineering - creating good features.
- Structured vs. unstructured - tabular data vs. raw text/images.
- Garbage in, garbage out - bad data produces bad results.
- Class imbalance - a rare class that makes plain accuracy misleading.
- Data sources - logs, email, network traffic, endpoints, user activity.

Review Questions (self-check for Quiz 2)

- In your own words, what is the difference between AI and machine learning?
- Give one task best solved with rules and one best solved with machine learning.
- Name three sources of security data and a question each can help answer.
- What is the difference between a feature and a label? Give an example.
- Why can a model be 99% accurate yet useless on imbalanced security data?
- Which chart would you use to compare counts of attack types, and which to track logins over time?

Bring your questions to class - and remember that AI tools are not permitted on quizzes and labs in this course.