EXAM 1 PRACTICE - ANSWER KEY

Practice Answer Key

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

Colorado State University - Computer Information Systems

About this document

Answers and brief explanations for the Exam 1 practice questions. Use this to self-check AFTER attempting the practice exam. Understanding WHY an answer is right matters more than the letter.

Part A - Multiple Choice: answers

1:C 2:B 3:A 4:D 5:A
6:C 7:D 8:A 9:B 10:C
11:A 12:B 13:C 14:D 15:A
16:D 17:A 18:B 19:B 20:C

Part A - explanations

1. C - Using a checksum to detect if a file was altered
Integrity = data is accurate/unaltered; a checksum detects tampering. (A is confidentiality, B is availability.)

2. B - Availability
A DDoS floods a system so it cannot serve users - an availability attack.

3. A - Vulnerability
A vulnerability is the weakness; the threat is the potential harm/actor; risk combines likelihood and impact.

4. D - Social engineering
Manipulating a person (not a machine) into giving access is social engineering.

5. A - Prevention
MFA blocks unauthorized access before it happens - a preventive control.

6. C - A hand-coded rule-based chatbot that follows fixed if/then rules
Rule-based systems follow written rules; they do not LEARN from data, so they are AI but not ML.

7. D - Label
The label is the answer we want to predict; features are the input clues.

8. A - Firewall and system logs
Logs are digital records a model can ingest; the others are not machine-readable data sources.

9. B - Relevant and good quality
'Garbage in, garbage out' - quantity helps only when the data is relevant and clean.

10. C - Organized in rows and columns like a table
Structured data is tabular - rows (examples) and columns (features).

11. A - Labeled examples with known answers
Supervised = learning from labeled examples (inputs paired with correct answers).

12. B - Estimate performance on data the model has not seen
A separate test set estimates real-world skill on unseen data, not memorization.

13. C - Overfitting
Memorizing training data and failing on new data is the definition of overfitting.

14. D - A yes/no question about a feature
Internal nodes ask feature-based yes/no questions; leaves give the prediction.

15. A - Smoother
Larger k averages over more neighbors, smoothing the boundary; very small k is noisy.

16. D - Regression
Predicting a number on a scale (dollars) is regression.

17. A - 70%
Accuracy = (TP + TN) / total = (40 + 30) / 100 = 70%.

18. B - False positive

A legitimate email wrongly flagged as spam is a false positive (a false alarm).

19. B - Rarely misses real positives

Recall = of the real positives, how many we caught; high recall = few misses. (Few false alarms = precision.)

20. C - Reduces false positives but increases false negatives

A stricter threshold flags less: fewer false alarms but more misses.

Part B - True / False: answers

21. FALSE

That describes availability; confidentiality is about keeping data secret from the wrong people.

22. TRUE

Phishing manipulates people into revealing information - classic social engineering.

23. FALSE

Rule-based software follows human-written rules; only machine learning learns from data.

24. TRUE

Supervised learning requires labeled examples (inputs paired with answers).

25. FALSE

Overfitting = performs well on training data but poorly on new data.

26. TRUE

k-NN takes the majority label of the closest known neighbors.

27. FALSE

Classification predicts a category; regression predicts a number.

28. TRUE

A false alarm: something legitimate flagged as positive/bad.

29. TRUE

Predicting the majority class scores high accuracy while catching none of the rare class.

30. FALSE

That is precision; recall is how many of the real positives were caught.

Part C - Short Answer: model answers

31. Explain the CIA triad and give a real example of violating each of the three goals.

Confidentiality (keep data secret) - e.g., a leaked customer database. Integrity (keep data accurate) - e.g., an attacker altering bank balances. Availability (keep systems usable) - e.g., ransomware locking files. Full credit: all three named + a valid violation example for each.

32. Give two differences between machine learning and rule-based software, and name one security task where ML is the better choice.

ML learns patterns from data vs. rules are written by a human; ML adapts/retrains as data changes vs. rules are fixed. Security task: detecting spam/phishing that constantly changes wording. Full credit: two clear differences + a sensible ML task.

33. What is overfitting, and how does a train/test split help you detect it?

Overfitting = the model memorizes the training data and performs poorly on new data. A held-out test set reveals it: high training accuracy but low test accuracy signals overfitting. Full credit: definition + how the split exposes it.

34. Explain the difference between a false positive and a false negative in fraud detection, and give the

business cost of each.

False positive = a legitimate transaction flagged as fraud (declined card, frustrated customer, support cost). False negative = real fraud allowed through (financial loss, chargebacks). Full credit: both defined correctly + a cost for each.

35. A model has 99% accuracy on a dataset where 99% of transactions are legitimate. Explain why this accuracy may be meaningless, and name a better metric.

With severe imbalance, a model that labels everything 'legitimate' scores 99% while catching zero fraud - accuracy hides this. Better: recall (catch rate of fraud) and/or precision. Full credit: imbalance explanation + a better metric named.

Part D - Applied Scenarios: model answers

36. A law firm's email filter (10 pts)

(a) False positive - legitimate email wrongly flagged.

(b) Precision is low: too many of the flagged items are actually legitimate.

(c) LOWER the threshold so it flags less aggressively (raising precision); trade-off: more spam may reach inboxes (more false negatives).

(d) e.g., a daily quarantine review, a 'not spam' report button, or allow-listing known client domains.

37. A hospital device-monitoring system (10 pts)

(a) Not well - supervised learning needs labeled attack examples, which they lack.

(b) Labeled records marking activity as 'attack' vs. 'normal'.

(c) Availability (devices/systems must be usable when needed).

(d) Unsupervised anomaly detection - it flags outliers without labels (previews Week 6).

38. A bank deploying a fraud model (10 pts)

(a) To estimate real performance on unseen transactions and avoid deploying an overfit model.

(b) Precision (of flagged, how many are real fraud - controls false alarms) and recall (of real fraud, how much is caught - controls misses).

(c) With 99% legitimate, a do-nothing model looks ~99% accurate while catching no fraud.

(d) Lower threshold = catch more fraud (higher recall) but more false alarms/declined cards; higher threshold = fewer false alarms but more missed fraud.