

Take-Home Test: Pattern Recognition and Anomaly Detection

Submission Date: 5 February, 2025

Instructions

- This is a take-home test. Answer all questions clearly and concisely.
- For numerical problems, show all calculations and steps.
- For case studies, provide detailed explanations with examples.
- Submit your answers in a neatly formatted document by the deadline.
- Late submissions will not be accepted.

Test Questions (3 Marks Each)

Theory-Based Questions

1. What is a pattern? Provide an example of a pattern in real-world data.
2. Define pattern recognition. Explain how it differs from anomaly detection.
3. List and briefly explain three pattern recognition techniques.
4. What is the difference between supervised and unsupervised learning in pattern recognition?
5. Explain the concept of overfitting in the context of pattern recognition.
6. What is anomaly detection? Provide an example of its application in cybersecurity.
7. How is anomaly detection calculated over time? Explain using a time-series example.
8. What is the role of probability theory in pattern recognition?
9. Explain the concept of maximum likelihood estimation (MLE) in pattern recognition.
10. What is a confusion matrix? How is it used to evaluate a classification model?

Numerical and Problem-Solving Questions

11. Given the following dataset, calculate the F1 score:
- True Positives (TP) = 50
 - False Positives (FP) = 10
 - False Negatives (FN) = 5
 - True Negatives (TN) = 100
12. Calculate the log-loss for the following predictions:
- True labels: [1, 0, 1, 1]
 - Predicted probabilities: [0.9, 0.2, 0.8, 0.6]
13. Given the following confusion matrix, calculate accuracy, precision, and recall:

	Predicted Yes	Predicted No
Actual Yes	80	20
Actual No	10	90

14. A polynomial regression model is fitted to the following data points:
- $X = [1, 2, 3, 4, 5]$
 - $Y = [2, 4, 9, 16, 25]$

Predict the value of Y when $X = 6$ using a polynomial of degree 2.

15. Using the following dataset, calculate the entropy:
- Class A: 30 instances
 - Class B: 20 instances
 - Class C: 10 instances

Case Study-Based Questions

16. Anomaly Detection in Healthcare
- A hospital uses an AI system to monitor patient vitals. The system flags unusual heart rates as anomalies. Explain how the system might use pattern recognition and anomaly detection to identify these anomalies.
17. Fraud Detection in Banking
- A bank uses anomaly detection to identify fraudulent transactions. Describe how the bank might use machine learning models to detect unusual spending patterns.
18. Predictive Maintenance in Manufacturing
- A manufacturing plant uses anomaly detection to predict equipment failures. Explain how the plant might use time-series data to detect anomalies and prevent downtime.

19. IBM Watson for Anomaly Detection

- IBM Watson is used for anomaly detection in network security. Describe how Watson might use machine learning to detect unusual network traffic patterns.

20. Document Analysis

- A company uses pattern recognition to classify documents into categories. Explain how the company might use supervised learning to train a model for document classification.

Additional Questions

21. What is the curse of dimensionality? How does it affect pattern recognition?
22. Explain the concept of hyperparameter tuning in machine learning.
23. What is the difference between precision and recall? Provide an example where precision is more important than recall.
24. How does information theory apply to feature selection in pattern recognition?
25. What is the role of regularization in preventing overfitting?

Submission Guidelines

- Submit your answers in a single PDF file. (Handwritten answers in notebook to be compiled)
- Name your file as: `YourName_EntryNo.pdf`.
- Email your submission to: *pooja.sarin@ddn.upes.ac.in* by **5 February, 2025, 11:59 PM**.
- Late submissions will not be accepted under any circumstances.

Marking Scheme

- Theory-Based Questions: 4 marks each (clarity, accuracy, and depth of explanation).
- Numerical Problems: 4 marks each (correct calculations and steps).
- Case Studies: 4 marks each (logical reasoning, application of concepts, and clarity).
- Additional Questions: 4 marks each

This take-home test is designed to evaluate your understanding of pattern recognition and anomaly detection concepts (Unit-1), as well as your ability to apply them to real-world scenarios. Good luck!