A4.3

Machine learning approaches

CHAPTER ASSIGNMENT · 32 MARKS

Name _____ Class _____ No. _____ Date _____

Answer every question in the space given. Show your working. You may use the chapter page to help you.

Questions from the Guru Jazzy chapter page, for classroom use.

Exam-style questions

1 A shop uses linear regression to predict spending. State its assumption, and describe how an outlier could affect it. [3]

2 An online retailer uses KNN to classify reviews. Outline how the choice of k affects results, and describe how to find a good k. [4]

3 A plant predicts equipment failure. Define precision and recall here, and explain why F1 may be better than accuracy. [4]

4 Outline how k-means clusters customers. [3]

5 Define support, confidence and lift for a library's borrowing data. [3]

6 A school uses a genetic algorithm to allocate electives. Describe a suitable fitness function, and outline selection and crossover. [4]

7 An energy company forecasts electricity demand using a neural network. Identify the problem type and suitable output layer, and define backpropagation. [4]

8 Describe the role of convolutional and pooling layers in a CNN. [4]

9 Outline cross-validation and why it matters. [3]

Chapter quiz

Q1 In y = a + bx, what does b represent?

A	B	C	D
The change in y for each increase of 1 in x	The value of y at x = 0	The error	The number of points

Q2 An R^2 near 1 means...

A	B	C	D
The line explains most of the variation	The model is useless	Slope is 1	There are no outliers

Q3 KNN is called a lazy learner because...

A	B	C	D
Most of its work happens at prediction time	It trains for a long time	It uses few features	It only predicts numbers

Q4 Why scale features before KNN?

A	B	C	D
Distances are sensitive to size	It removes labels	It adds neighbours	It is never needed

Q5 For cancer screening, which metric matters most?

A	B	C	D
Recall	Precision	Training speed	Depth

Q6 Which is a hyperparameter?

A	B	C	D
k in KNN	A weight learned in training	A label	A row of data

Q7 Good on training data and poor on test data suggests...

A	B	C	D
Overfitting	Underfitting	Perfect fit	Clustering

Q8 Which clustering method labels noise?

A	B	C	D
DBSCAN	K-means	Linear regression	KNN

Q9 A lift above 1 means...

A	B	C	D
A makes B more likely than B's usual rate	A prevents B	A and B are independent	B is rare

Q10 In Q-learning, ε controls...

A	B	C	D
How often the agent explores	The size of each update	How much the future counts	The number of states

Q11 A genetic algorithm's fitness function...

A	B	C	D
Scores each candidate solution	Swaps two genes	Ends the program	Labels the data

Q12 What does mutation protect against?

A	B	C	D
Loss of diversity	Slow networks	Missing labels	Large files

Q13 Which layer gives a CNN its smaller feature maps?

A	B	C	D
Pooling	Input	Softmax	Bias

Q14 Why add an activation function after convolution?

A	B	C	D
To add non-linearity	To add pixels	To encrypt the image	To remove the filters