



**Introduction to
Machine Learning**

Assignment- Week 1

TYPE OF QUESTION: MCQ

Number of questions: 10

Total mark: 10 X 2 = 20

MCQ Question

QUESTION 1:

Which of the following are classification tasks?

- A. Detect pneumonia from chest X-ray image
- B. Predict the price of a house based on floor area, number of rooms etc.
- C. Predict the temperature for the next day
- D. Predict the amount of rainfall

Correct Answer: A

Detailed Solution : The number of classes in pneumonia detection is discrete. So, it's a classification task. In other options, the output variable is a continuous class, so these are regression tasks.

QUESTION 2:

Which of the following is not a type of supervised learning?

- A. Classification
- B. Regression
- C. Clustering
- D. None of the above

Correct Answer: C. Clustering

Detailed Solution : Classification and Regression are both supervised learning methods as they need class labels or target values for training, but Clustering doesn't need target values.



QUESTION 3:

Which of the following tasks is NOT a suitable machine learning task?

- A. Finding the shortest path between a pair of nodes in a graph
- B. Predicting if a stock price will rise or fall
- C. Predicting the price of petroleum
- D. Grouping mails as spams or non-spams

Correct Answer : A. Finding the shortest path between a pair of nodes in a graph

Detailed Solution : Finding the shortest path is a graph theory based task, whereas other options are completely suitable for machine learning.

QUESTION 4:

Suppose I have 10,000 emails in my mailbox out of which 300 are spams. The spam detection system detects 150 mails as spams, out of which 50 are actually spams. What is the precision and recall of my spam detection system?

- A. Precision = 33.33%, Recall = 25%
- B. Precision = 25%, Recall = 33.33%
- C. Precision = 33.33%, Recall = 16.66%
- D. Precision = 75%, Recall = 33.33%

Correct Answer: C

Detailed Solution :

$$\begin{aligned}\text{Precision} &= \frac{Tp}{Tp+Fp} \\ &= \frac{50}{50+100} \\ &= 33.33\%\end{aligned}$$

$$\begin{aligned}\text{Recall} &= \frac{Tp}{Tp+Fn} \\ &= \frac{50}{50+250} \\ &= 16.66\%\end{aligned}$$



QUESTION 5 :

Which of the following is/are supervised learning problems?

- A. Predicting disease from blood samples.
- B. Grouping students in the same class based on similar features.
- C. Face recognition to unlock your phone.

Correct Answer: A, C

Detailed Solution: Option B is an unsupervised learning problem.

QUESTION 6:

Aliens challenge you to a complex game that no human has seen before. They give you time to learn the game and develop strategies before the final showdown. You choose to use machine learning because an intelligent machine is your only hope. Which machine learning paradigm should you choose for this?

- A. Supervised learning
- B. Unsupervised learning
- C. Reinforcement learning
- D. Use a random number generator and hope for the best

Correct Answer: C. Reinforcement learning

Detailed Solution: Reinforcement learning is the optimal method for building agents for complex games where no expert trajectories exist. It is possible to design a reward function/signal that depends on the outcome of the game. The objective of the player agent is to maximize the total reward collected from the game.



QUESTION 7:

How many Boolean functions are possible with N features?

- A. (2^{2^N})
- B. (2^N)
- C. (N^2)
- D. (4^N)

Correct Answer: A. (2^{2^N})

Detailed Solution : There are 2^N possible combinations of N input boolean features. Each of these input feature vectors can be either True (1) or False (0), so there can be 2^{2^N} possible truth tables for the boolean function.

QUESTION 8:

What is the use of Validation dataset in Machine Learning?

- A. To train the machine learning model.
- B. To evaluate the performance of the machine learning model
- C. To tune the hyperparameters of the machine learning model
- D. None of the above.

Correct Answer : C. To tune the hyperparameters of the machine learning model

Detailed Solution : The validation dataset is used to tune the model's hyperparameters during training



QUESTION 9:

Regarding bias and variance, which of the following statements are true? (Here 'high' and 'low' are relative to the ideal model.)

- A. Models which overfit have a high bias.
- B. Models which overfit have a low bias.
- C. Models which underfit have a high variance.
- D. Models which underfit have a low variance.

Correct Answer : B, D

Detailed Solution : In supervised learning, underfitting happens when a model is unable to capture the underlying pattern of the data. These models usually have high bias and low variance. Overfitting happens when our model captures the noise along with the underlying pattern in data. These models have low bias and high variance.

QUESTION 10:

Which of the following is a categorical feature?

- A. Height of a person
- B. Price of petroleum
- C. Mother tongue of a person
- D. Amount of rainfall in a day

Correct Answer: C

Detailed Solution: Categorical variables represent types of data which may be divided into groups. Mother tongue is a categorical feature. All other options are continuous.

*****END*****

