Recommender Systems

Assignment- Week 2

TYPE OF QUESTION: MCQ/MSQ

Number of questions: 10 **Total marks: 2×10**

QUESTION 1:

Which of the following is a not task under data preprocessing?

a) Data discretization
b) Data integration
c) Data transformation
d) Data visualization

Correct Answer: d

Explanation: Refer to Week 2 lecture 1 slide 4

QUESTION 2:

For multi variable data, which of the following is not a right method for outlier detection.

a) Box plot
b) Mahalnobis Distance
c) Distance based measures
d) One class classifiers

Correct Answer: a

Explanation: Refer to Week 2 lecture 1 slide 9

QUESTION 3:

If **M** is a square matrix, **λ** is one of its eigen value and **e** is the corresponding eigen vector. Let **c** be any constant. Based on this which of the following statement is true

a) c**λ** is an eigenvector of M with e as the eigenvalue
b) c**λ** is an eigenvalue of M with **e** as an eigenvector
c) ce is an eigenvector of M with the same eigenvalue
d) **ce** is an eigenvector of M with c**λ** as the eigenvalue

Correct Answer: c

Explanation: Refer to Week 2 lecture 2 slide 5

QUESTION 4:

We go for Singular Value Decomposition (SVD) of a matrix M=UΣV. The singular values are in matrix Σ. Which of following is true for Σ

a) It is a square matrix
b) The off diagonal elements are all zero
c) It is a r×r matrix, where r is the rank of M
d) **All of the above**.

Correct Answer: d

Explanation: Refer to Week 2 lecture 2 slide 14

QUESTION 5:

Both regression type models and classification model have the following common characteristics

a) Both are supervised and require example patterns for learning
b) Both are supervised and do not require example patterns for learning
c) Both are supervised and require example patterns for learning
d) Both are supervised and require example patterns for learning

Correct Answer: a

Explanation: Refer to Week 2 lecture 3 slide 5-8

QUESTION 6:

Both Ridge regression and Lasso have penalty term in the error minimization problem. The difference is

a) Ridge has l_2 penalty whereas Lasso has both l_1 and l_2 penalty
b) Ridge has l_1 penalty whereas Lasso has l_2 penalty
c) Ridge has l_2 penalty whereas Lasso has l_1 penalty
d) Ridge has both l_1 and l_2 penalty whereas Lasso has l_1 penalty

Correct Answer: c

Explanation: Refer to Week 2 lecture 3 slide 11-12

QUESTION 7:

What is the primary concept behind the K-Nearest Neighbors (KNN) algorithm?

a) Using a decision tree to classify data points based on features
b) **Identifying the K-closest data points in the feature space and making predictions based on their labels**.
c) Calculating the gradient of a cost function to update parameters iteratively.
d) Clustering data into K groups based on their similarity

Correct Answer: b

Explanation: Refer to Week 2 lecture 4 slide 4-5

QUESTION 8:

What is a key difference between Naive Bayes classifier and Bayesian Networks?

a) Naive Bayes assumes conditional independence of features, while Bayesian Networks model dependencies between features explicitly.
b) Naive Bayes can represent complex dependencies among variables, unlike Bayesian Networks.
c) Bayesian Networks require a larger dataset compared to Naive Bayes for effective training.
d) Naive Bayes is used for regression tasks, while Bayesian Networks are exclusively for classification tasks.

Correct Answer: a

Explanation: Refer to Week 2 lecture 4 slide 9-10

QUESTION 9:

What is the primary objective of the K-Means clustering algorithm?

a) To minimize the sum of squared distances between each data point and its assigned cluster centroid.
b) To maximize the number of clusters for better granularity.
c) To assign each data point to multiple clusters simultaneously.
d) To create hierarchical clusters based on data point density.

Correct Answer: a

Explanation: Refer to Week 2 lecture 5 slide 10-14

QUESTION 10:

Which of the following is a limitation of K-Means clustering?

a) It can only be used for supervised learning tasks.
b) The number of clusters *k* must be specified beforehand.
c) It cannot handle non-linear relationships between data points effectively.
d) It cannot guarantee finding the global minimum of the objective function.

Correct Answer: b

Explanation: Refer to Week 2 lecture 5 slide 10-14

**********END**********