MACHINE LEARNING FOR SOIL AND CROP MANAGEMENT

Assignment- Week 3

TYPE OF QUESTION: MCQ/MSQ

Number of questions: 15 **Total mark: 15 X 1 = 15**

QUESTION 1:

Principal components analysis (PCA) is a/an_____approach for deriving a _____ dimensional set of features from large set of correlated variables.

a. Supervised, high
b. Unsupervised, low
c. Supervised, low
d. Unsupervised, high

Correct Answer: b

Detailed Solution: PCA is an unsupervised approach for deriving low-dimensional set of features (called principal components) from large set of correlated variables. Principal components allow us to summarize this set with a smaller number of representative variables that collectively explain most of the variability in the original dataset.

QUESTION 2:

In PCA, the eigen vectors represent-

a. Direction of maximum variance
b. Eliminated features
c. Original variables
d. Noise in the data

Correct Answer: a

Detailed Solution: In PCA, eigen vectors represent the direction of maximum variance in the data. They define the axes or principal components of the new feature space.

QUESTION 3:

What is the purpose/use of screeplot in PCA?

a. To eliminate features
b. To determine how many principal components to retain
c. To scale the data
d. Visualize correlation between variables

Correct Answer: b

Detailed Solution: Screeplot helps to visualize the proportion of variance explained by each principal component to decide how many components to retain.

QUESTION 4:

In PCA, what does PC1 represent?

a. The least variability direction
b. The direction of greatest variability
c. The average of all features
d. A threshold value for variance

Correct answer: b

Detailed Explanation: PC1 captures the largest variance in the data, making it the direction of greatest variability.

QUESTION 5:

In CART the feature space is partitioned into a set of _____, and fit a simple model in each leaf.

a. Square
b. Circle

c. Triangle

d. Rectangle

Correct Answer: d

Detailed Solution: In classification and regression trees (CART) tree partitions the feature space into a set of rectangles, and fit a simple model in each leaf (terminal node).

QUESTION 6:

Which of the following statement is correct regarding the PCA biplot?

a. When two vectors are close, forming a small angle, the two variables they represent are positively correlated.

b. If the two vectors meet each other at 90°, they are not likely to be correlated.

c. When the two vectors diverge and form a large angle, they are negatively correlated.

d. All are correct

Correct Answer: d

Detailed Solution: When two vectors are close, forming a small angle, the two variables they represent are positively correlated. If they meet each other at 90°, they are not likely to be correlated. And, if the two vectors diverge and form a large angle, they are negatively correlated.

QUESTION 7:

What is the key output of PLSR?

a. Hyperplane

b. Principal components

c. Latent factors

d. Gini index value

Correct Answer: c

Detailed Solution: PLSR generates latent factors that maximize covariance.

QUESTION 8:

The main purpose of pruning in CART is-

a. Simplifying the tree and prevent overfitting
b. Increasing impurity
c. Adding more nodes to the tree
d. Combining features

Correct Answer: a

Detailed Solution: Pruning reduces the overfitting by removing unnecessary branches.

QUESTION 9:

In CART, which impurity measure is used to determine the best split?

a. Eigen values
b. Principal components
c. Regression coefficient
d. Gini index

Correct Answer: d

Detailed Solution: Gini index measures impurity and helps in identifying the best split.

QUESTION 10:

PCA transfers a set of _____ variables into a new set of _____ variables.

a. Correlated, correlated
b. Correlated, uncorrelated
c. Uncorrelated, correlated
d. Uncorrelated, uncorrelated

Correct Answer: b

Detailed Solution: PCA transfers a set of correlated variables into a new set of uncorrelated variables.

QUESTION 11:

What is the key difference between PCR and Partial Least Squares Regression (PLSR)?

a. PCR focuses on X; PLSR focuses on X and Y
b. PCR is more complex than PLSR
c. PCR eliminates variables; PLSR combines variables
d. There is no difference

Correct answer: a

Detailed explanation: PCR uses principal components that explain variance in X, while PLSR selects components that explain covariance between X and Y.

QUESTION 12:

Which algorithm constructs decision trees using bootstrap samples?

a. Principal Component Regression

b. Random Forest

c. Support Vector Machine

d. Partial Least Squares Regression

Correct Answer: b

Detailed Solution: Random Forest uses bootstrap aggregation (bagging) to construct decision trees.

QUESTION 13:

Dimensionality reduction involves?

a. Feature elimination
b. Feature addition
c. Feature extraction
d. Both a and c

Correct Answer: d

Detailed Solution: Both feature elimination and feature extraction are involved in Dimensionality reduction.

QUESTION 14:

In support vector machine (SVM), _____ the margin, _____ the generalization error of the classifier.

a. Smaller, lower
b. Larger, lower
c. Larger, higher
d. None of these

Correct Answer: b

Detailed Solution: In support vector machine (SVM), the larger the margin, the lower the generalization error of the classifier.

QUESTION 15:

Which technique uses a hyperplane for classification?

a. Random Forest
b. Partial Least Square Regression
c. Support Vector Machine
d. Principal Component Analysis

Correct Answer: c

Detailed Solution: SVM constructs a hyperplane to classify data into distinct classes.