MACHINE LEARNING FOR SOIL AND CROP MANAGEMENT

Assignment- Week 11

TYPE OF QUESTION: MCQ/MSQ

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

QUESTION 1:

_____ is an organization that maintains a geodetic parameter database with standard codes.

    a. EPSG

    b. WSI

    c. IWMI

    d. SRS

Correct Answer: a

Detailed Solution: EPSG stands for European Petroleum Survey Group is an organization that maintains a geodetic parameter database with standard codes.

QUESTION 2:

_____ is a storage format that contains the coordinate reference system represented as text.

    a. .shp

    b. .dbf

    c. .shx

    d. .prj

Correct Answer: d

Detailed Solution: .prj is a storage format that contains the coordinate reference system represented as text.

QUESTION 3:

A _____ is an Esri vector data storage format for storing the location, shape, and attributes of geographic features. It is stored as a set of related files and contains one feature class.

    a. Folder

    b. File

    c. Shapefile

    d. Database

Correct Answer: c

Detailed Solution: A shapefile is an Esri vector data storage format for storing the location, shape, and attributes of geographic features. It is stored as a set of related files and contains one feature class.

QUESTION 4:

The package _____ is a collection of algorithms related to modelling of soil resources, soil classification, soil profile aggregation, and visualization.

    a. ggplot2

    b. aqp

    c. sp

    d. RSAGA

Correct Answer: b

Detailed Solution: The package aqp is a collection of algorithms related to modelling of soil resources, soil classification, soil profile aggregation, and visualization.

QUESTION 5:

_____ package is an implementation of the grammar of graphics in R. It combines the advantages of both base and lattice graphics: conditioning and shared axes are handled automatically, and you can still build up a plot step by step from multiple data sources.

    a. Gam

    b. Gstat

    c. caret

    d. ggplot2

Correct Answer: d

Detailed Solution: The ggplot2 package is an implementation of the grammar of graphics in R. It combines the advantages of both base and lattice graphics: conditioning and shared axes are handled automatically, and you can still build up a plot step by step from multiple data sources.

QUESTION 6:

The traditional method of sampling soil involves dividing a soil profile into _____ based on the attributes easily observed in the field, such as morphological soil properties.

    a. Solum

    b. Horizon

    c. Pedon

    d. None of the above

Correct Answer: b

Detailed Solution: The traditional method of sampling soil involves dividing a soil profile into horizon based on the attributes easily observed in the field, such as morphological soil properties.

QUESTION 7:

Which soil horizon is typically observed at the top surface of a soil profile, especially in forests, and consists of organic matter such as decomposed leaves and plant material?

    a. A Horizon

    b. B Horizon

    c. O Horizon

    d. C Horizon

Correct Answer: c

Detailed Solution: The O horizon is typically observed at the top surface of a soil profile, especially in forests, and consists of organic matter such as decomposed leaves and plant material.

QUESTION 8:

Which of the following statement is true in case of DSM

    a. Deriving a continuous function using the available horizon data as some input.

    b. Preserving original data so that it can be retrieved again via integration of the continuous spline

    c. Continuous depth function like the equal-area quadratic spline function is used

    d. All of the above

Correct Answer: d

Detailed Solution: The solution to use the discontinuous legacy data in DSM is to derive a continuous function using the available horizon data as some input. Continuous depth function like the equal-area quadratic spline function is used. Further, original data is preserved so that it can be retrieved again via integration of the continuous spline.

QUESTION 9:

What is the full form of DSM?

    a. Digital Soil Model

    b. Direct Soil Mapping

    c. Digital Soil Mapping

    d. None of the above

Correct Answer: c

Detailed Solution: The full form of DSM is digital soil mapping.

QUESTION 10:

Which of the following skewness value represent a highly skewed distribution?

    a. > +1 or < -1

    b. +0.5 to +1

    c. – 0.5 to + 0.5

    d. None of the above

Correct Answer: a

Detailed Solution: The skewness value > +1 or < -1 indicates that the distribution is highly skewed

QUESTION 11:

The sharpness of the peak of a frequency-distribution curve is known as _____

    a.   Skewness
    b.   Normality
    c.   Kurtosis
    d.   None of the above

Correct Answer: c

Detailed Solution: The sharpness of the peak of a frequency-distribution curve is known as Kurtosis.

QUESTION 12:

Which of the following statement is true?

    a.   Skewness is the distortion or asymmetry that deviates from the symmetrical bell curve, or normal distribution, in a set of data.
    b.   Kurtosis identifies whether the tails of a given distribution contain extreme value.
    c.   If the kurtosis value is greater than +1, the distribution is too peaked.
    d.   All of the above

Correct Answer: d

Detailed Solution: Skewness is the distortion or asymmetry that deviates from the symmetrical bell curve, or normal distribution, in a set of data. Kurtosis identifies whether the tails of a given distribution contain extreme value. a. If the kurtosis value is greater than +1, the distribution is too peaked.

QUESTION 13:

_____is a single statistic that both evaluates the accuracy and precision of the relationship. It is often referred to as the goodness of fit along a 45-degree line.

   a.   Bias
   b.   Lin's concordance correlation coefficient
   c.   Mean square error
   d.   Regression coefficient

Correct Answer: b

Detailed Solution: Lin's concordance correlation coefficient is a single statistic that both evaluates the accuracy and precision of the relationship. It is often referred to as the goodness of fit along a 45-degree line.

QUESTION 14:

CUBIST model is popular in DSM community because

   a. It can mine non-linear relationship
   b. It does not have finite prediction issue
   c. Both a and b
   d. None of these

Correct Answer: c

Detailed Solution: Cubist model is popular in DSM community. Its popularity is due to its ability to "mine" non-linear relationships in data, but does not have the issues of finite predictions that occur for other decision and regression tree models

QUESTION 15:

Which of the following statement is true for the CUBIST model?

a. Based on the M5 algorithm of Quinlan(1992)

b. The Cubist model first partitions the data into subsets within which their characteristics are similar

c. It apply regression to predict the soil property if conditions met, else apply the next rule

d. All of the above

Correct Answer: d

Detailed Solution: The CUBIST model is based on the M5 algorithm of Quinlan(1992). The Cubist model first partitions the data into subsets within which their characteristics are similar. It applies regression to predict the soil property if conditions met, else apply the next rule