

<https://swayam.gov.in>https://swayam.gov.in/nc_details/NPTEL

manojmanu609bv@gmail.com ▾

NPTEL (<https://swayam.gov.in/explorer?ncCode=NPTEL>) » Data Science for Engineers (course)Course
outlineAbout
NPTEL ()How does an
NPTEL
online
course
work? ()Setup Guide
()Pre Course
Material ()

Week 0 ()

Week 1 ()

Week 2 ()

Week 3 ()

Week 4 ()

Week 5 ()

Week 6 ()

Week 7 : Assignment 7

The due date for submitting this assignment has passed.

Due on 2024-09-11, 23:59 IST.

Assignment submitted on 2024-09-09, 10:54 IST

1) Which among the following is not a type of cross-validation technique?

1 point

- ☐ LOOCV
- ☐ k-fold cross validation
- ☐ Validation set approach
- ☒ Bias variance trade off

Yes, the answer is correct.

Score: 1

Accepted Answers:

Bias variance trade off

2) Which among the following is a classification problem?

1 point

- ☐ Predicting the average rainfall in a given month.
- ☒ Predicting whether a patient is diagnosed with a disease or not.
- ☐ Predicting the price of a house.
- ☒ Predicting whether it will rain or not tomorrow.

Yes, the answer is correct.

Score: 1

Accepted Answers:

*Predicting whether a patient is diagnosed with a disease or not.**Predicting whether it will rain or not tomorrow.*

Week 7 ()

- ☐ Cross Validation (unit? unit=85&lesson=86)
- ☐ Multiple Linear Regression Modelling Building and Selection (unit? unit=85&lesson=87)
- ☐ Classification (unit? unit=85&lesson=88)
- ☐ Logistic Regression (unit? unit=85&lesson=89)
- ☐ Logistic Regression (Continued) (unit? unit=85&lesson=90)
- ☐ Performance Measures (unit? unit=85&lesson=91)
- ☐ Logistic Regression Implementation in R (unit? unit=85&lesson=92)
- ☐ Dataset (unit? unit=85&lesson=93)
- ☐ FAQ (unit? unit=85&lesson=94)
- ☐ Practice: Week 7: Assignment 7

Consider the following confusion matrix for the classification of Hatchback and SUV:

		True	
		Hatchback	SUV
Prediction	Hatchback	55	5
	SUV	0	40

3) Find the accuracy of the model.

1 point

- ☒ 0.95
- ☐ 0.55
- ☐ 0.45
- ☐ 0.88

Yes, the answer is correct.

Score: 1

Accepted Answers:

0.95

4) Find the sensitivity of the model.

1 point

- ☐ 0.95
- ☐ 0.55
- ☒ 1
- ☐ 0.88

Yes, the answer is correct.

Score: 1

Accepted Answers:

1

5) Under the 'family' parameter of glm() function, which one of the following distributions correspond to logistic regression for a variable with binary output?

1 point

- ☒ Binomial
- ☐ Gaussian
- ☐ Gamma
- ☐ Poisson

Yes, the answer is correct.

Score: 1

Accepted Answers:

Binomial

Use the following information to answer Q6, Q7, Q8, Q9, and Q10:

Load the dataset **iris.csv**

(https://drive.google.com/file/d/1SqGPxYa8xCDBc5ZsLN0tCzUceoVmR_6e/view?usp=sharing) as a dataframe irisdata, with the first column as index headers, first row as column headers, dependent variable as factor variable, and answer the following questions.

(Non Graded)
(assessment?
name=212)

● **Quiz: Week 7
: Assignment
7**
(assessment?
name=223)

○ Week 7
Feedback
Form : Data
Science for
Engineers
(unit?
unit=85&lesso
n=159)

Week 8 ()

**Text
Transcripts
()**

**Download
Videos ()**

Books ()

**Problem
Solving
Session -
July 2024 ()**

The iris dataset contains four Sepal and Petal features (Sepal Length, Sepal Width, Petal Length, Petal Width, all in cm) of 50 equal samples of 3 different species of the iris flower (Setosa, Versicolor, and Virginica).

6) What is the dimension of the dataframe?

1 point

- ☒ (150, 5)
☐ (150, 4)
☐ (50, 5)
☐ None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:
(150, 5)

7) What can you comment on the distribution of the independent variables in the dataframe?

1 point

- ☐ The variables Sepal Length and Sepal Width are not normally distributed
☒ All the variables are normally distributed
☐ The variable Petal Length alone is normally distributed
☐ None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:
All the variables are normally distributed

8) How many rows in the dataset contain missing values?

1 point

- ☐ 10
☐ 5
☐ 25
☒ 0

Yes, the answer is correct.

Score: 1

Accepted Answers:
0

9) Which of the following code blocks can be used to summarize the data (finding the mean of the columns PetalLength and PetalWidth), similar to the one given below. **1 point**



PetalLength	PetalWidth
3.758000	1.199333

- ☒ lapply(irisdata[, 3:4], mean)
☐ sapply(irisdata[, 3:4], 2, mean)
☒ apply(irisdata[, 3:4], 2, mean)
☐ apply(irisdata[, 3:4], 1, mean)

Yes, the answer is correct.

Score: 1

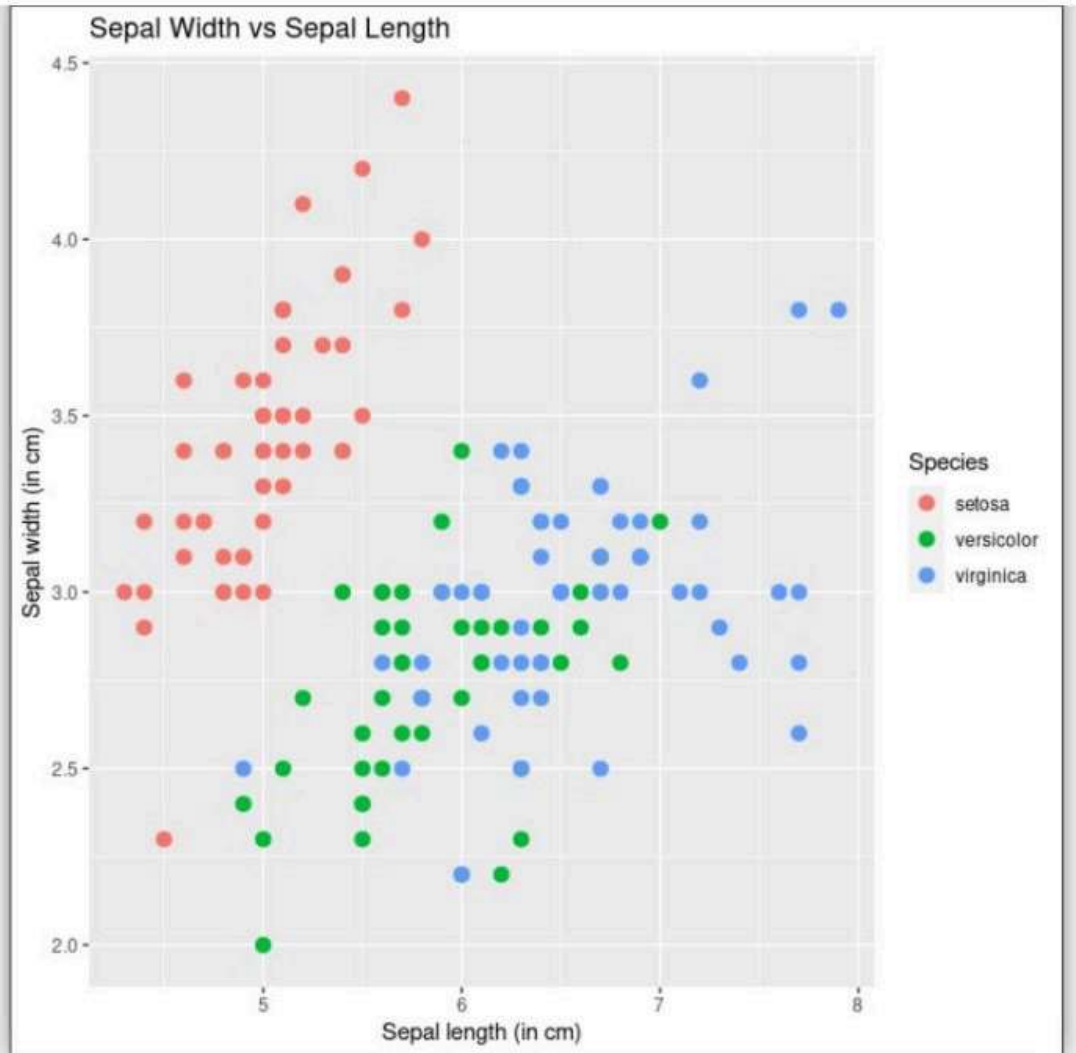
Accepted Answers:

`lapply(irisdata[, 3:4], mean)`

`apply(irisdata[, 3:4], 2, mean)`

10) What can be interpreted from the plot shown below?

1 point



- ☐ Sepal widths of Versicolor flowers are lesser than 3 cm.
- ☒ Sepal lengths of Setosa flowers are lesser than 6 cm.
- ☐ Sepal lengths of Virginica flowers are greater than 6 cm.
- ☒ Sepals of Setosa flowers are relatively more wider than Versicolor flowers.

Yes, the answer is correct.

Score: 1

Accepted Answers:

Sepal lengths of Setosa flowers are lesser than 6 cm.

Sepals of Setosa flowers are relatively more wider than Versicolor flowers.