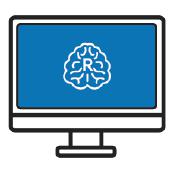
Data Analytics



SESSION 16: Association

Analysis

Assignment 1

1

Data Analytics

Table of Contents

1.Introduction 3

2.Objective 3

3.Prerequisites 3

4.Associated Data Files 3

5.Problem Statement 3

6.Expected Output 3

7.Approximate Time to Complete Task 3

2

Data Analytics

**1. Introduction**

This assignment will help you understand the concepts learnt in the session.

**2. Objective**

This assignment will test your skills on the Association Analysis Algorithm using R.

**3. Prerequisites**

Not applicable.

**4. Associated Data Files**

Not applicable.

**5. Problem Statement**

1. Use the below given data set

Data Set

2. Perform the below given activities:

a. Predict the no of comments in next H hrs

Note:-

1. Use LASSO, Elastic Net and Ridge and other regression techniques that are covered in the module
2. Report the training accuracy and test accuracy
3. compare with linear models and report the accuracy
4. create a graph displaying the accuracy of all models

|  |
| --- |
| #Problem |
|  | #1. Use the below given data set |
|  | #Data Set |
|  | #2. Perform the below given activities: |
|  | # a. Predict the no of comments in next H hrs |
|  | #Note:- |
|  | # 1. Use LASSO, Elastic Net and Ridge and other regression techniques that are covered in the |
|  | # module |
|  | #2. Report the training accuracy and test accuracy |
|  | #3. compare with linear models and report the accuracy |
|  | #4. create a graph displaying the accuracy of all models |
|  |  |
|  | #Answers |
|  | #a),b),c),d) |
|  | #using slr dataset |
|  | #reading the dataset and viewing |
|  | slr <- read.csv("D:/acadgild/slr.csv") |
|  | slr1<- slr |
|  | View(slr1) |
|  |  |
|  | #features |
|  | dim(slr1) |
|  | str(slr1) |
|  |  |
|  | library(psych) |
|  | describe(slr1) |
|  | summary(slr1) |
|  |  |
|  | #visualization |
|  | hist(slr1$Advt ,xlab = "advt", ylab = "Frequency",main="Histogram of advt",col="red") |
|  | hist(slr1$Sales ,xlab = "sales", ylab = "Frequency",main="Histogram of sales",col="blue") |
|  |  |
|  | plot(slr1$Advt,slr1$Sales) |
|  |  |
|  | #\*\*\*NOTE\*\*\* |
|  | #using linear regression model technique |
|  | #using slr1 dataset |
|  | #linear regression model |
|  | model<- lm(slr1$Advt~slr1$Sales) |
|  | model |
|  |  |
|  | #Important features |
|  |  |
|  | #multiple r squared value |
|  | #p value of slope test |
|  | #F stats |

3

Data Analytics

**6. Expected Output**

N/A

**7. Approximate Time to Complete Task**

30 mins.

4