Experiment -6

1. Use Principle component analysis for the given data set and find the better prediction model for quantitative data.

2. For the given data, based on the scree plot, find out the number of principle components and in principle component analysis.

3. Plot Principle Component Analysis output and explain the visualization part of the dimension reduction techniques.

Program:

```
# Load Packages and prepare dataset

library(TH.data)

library(caret)

data("GlaucomaM", package = "TH.data")

trainData <- GlaucomaM

View(trainData)

dim(trainData)

head(trainData)

str(trainData)

set.seed(100)

options(warn = -1)


subsets <- c(1:5, 10, 15, 18)


ctrl <- rfeControl(functions = rfFuncs,

           method = "repeatedcv",

           repeats = 5,

           verbose = FALSE)
```

```
lmProfile <- rfe(x = trainData[, c(1:3, 5:13)], y=trainData$ozone_reading,
        sizes = subsets,
        rfeControl = ctrl)

lmProfile

input = read.csv("H:\\JGi Classes\\Dimentionality Reduction and Model Validation VII Sem 2018-22\\Lab Programs\\data files\\iris.csv")
names(input)
str(input)

model = prcomp(input[,1:4], scale=TRUE)
model$sdev
model$rotation
model$center
model$scale

par(mfrow=c(2,2))
plot(model$x[,1], col=input[,5])
plot(model$x[,2], col=input[,5])
plot(model$x[,3], col=input[,5])
plot(model$x[,4], col=input[,5])

model$sdev^2 / sum(model$sdev^2)
plot(model)
```

## PCA without 'prcomp'

## Normalize the input feature.

```
input$sepal_len1 = (input$sepal_len - mean(input$sepal_len) )/sd(input$sepal_len)

input$sepal_wid1 = (input$sepal_wid - mean(input$sepal_wid))/sd(input$sepal_wid)

input$petal_len1 = (input$petal_len - mean(input$petal_len))/sd(input$petal_len)

input$petal_wid1 = (input$petal_wid - mean(input$petal_wid))/sd(input$petal_wid)
```

##Get the covarience matrix and eigen vector.

```
matrix_form = matrix(c(input$sepal_len1, input$sepal_wid1, input$petal_len1, input$petal_wid1),
ncol=4)

m = cov(matrix_form)

eigenV = eigen(m)

eigenV$vectors
```

Output: