Date: 10/April/2024
EXPERIMENT – 10

K- NEAREST NEIGBOR CLASSIFIER

AIM: To perform K nearest neighbor classifier

SOFTWARE REQUIRED: RStudio

R CODE:

```
rm(list=ls())
data <- read.csv ("CreditWorthiness (1).csv", stringsAsFactors =</pre>
               line
                     reads the
       #This
                                  data from the
                                                     CSV
                                                            file
"CreditWorthiness.csv" into a data frame called data, with strings
converted to factors.
str(data)
summary(data)
plot(data)
data$Cdur <-as.integer(data$Cdur)</pre>
data$Cpur <-as.integer(data$Cpur)</pre>
data$Camt <- as.integer(data$Camt)</pre>
data$age <- as.integer(data$age)</pre>
data[, -5] <- scale(data[, -5])</pre>
set.seed(123)
train indices <- sample (nrow (data), 900)
data TRAIN <- data[train indices, ]</pre>
data TEST <- data[-train indices, ]</pre>
library (class)
library (caret)
knnpredict <- knn(train = data TRAIN[, -5], test= data TEST[,- 5],</pre>
cl = data TRAIN$creditScore, k = 5)
confusionMatrix(table(knnpredict, data TEST$creditScore), positive
= 'good')
```

OUTPUT:

BCSE352E–Essentials of Data Analytics – Lab [Winter Semester 2023–24]

Name of the Student: **Atharv Patel** Register Number: **21BLC1560**

e. > summary(data) Cdur Cpur creditScore Camt age Min. : 4.0 electronics :280 Min. : 2380 Min. :19.00 bad :300 1st Qu.:12.0 second hand vehicle:234 1st Qu.: 13535 1st Qu.:27.00 good:700 Median :18.0 furniture :181 Median: 23075 Median: 33.00 Mean :20.9 new vehicle :103 Mean : 32593 Mean :35.55 : 97 3rd Qu.:24.0 Business 3rd Qu.: 39603 3rd Qu.:42.00 Max. :72.0 education : 50 Max. :184120 Max. :75.00 (Other) : 55

> contrustonmacrix(capte(knipredict, data_test

Confusion Matrix and Statistics

knnpredict bad good bad 9 12 58 good 21

Accuracy: 0.67

95% CI: (0.5688, 0.7608)

No Information Rate: 0.7 P-Value [Acc > NIR] : 0.7793

Kappa : 0.1406

Mcnemar's Test P-Value: 0.1637

Sensitivity: 0.8286 Specificity: 0.3000 Pos Pred Value: 0.7342 Neg Pred Value: 0.4286 Prevalence: 0.7000 Detection Rate: 0.5800

Detection Prevalence: 0.7900

BCSE352E–Essentials of Data Analytics – Lab [Winter Semester 2023–24]

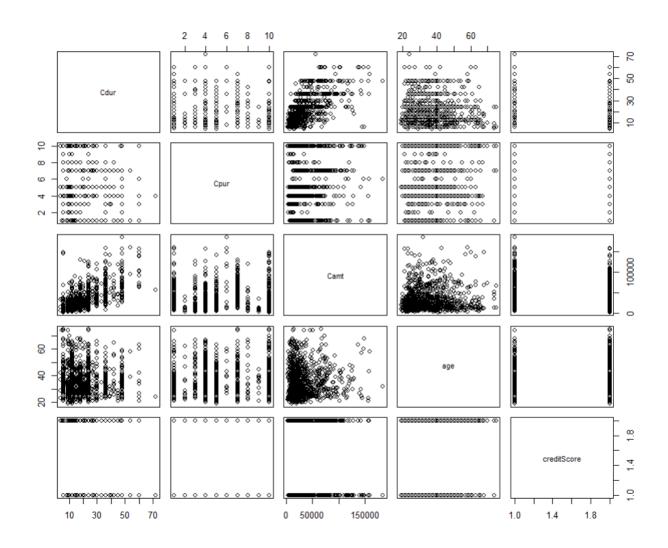
Name of the Student: Atharv Patel Register Number: 21BLC1560

Mcnemar's Test P-Value: 0.1637

Sensitivity: 0.8286 Specificity: 0.3000 Pos Pred Value: 0.7342 Neg Pred Value: 0.4286 Prevalence: 0.7000 Detection Rate: 0.5800

Detection Prevalence: 0.7900 Balanced Accuracy: 0.5643

'Positive' Class : good



Name of the Student: **Atharv Patel** Register Number: **21BLC1560**