#Project : HW_09_SVM #Purpose : SVM methodology on breast-cancer-wisconsin dataset #First Name : Omkar #Last Name : Sinha #CWID : 10468312 #Date : 12/06/2021

```r
rm(list=ls())
cat("\014")
```

```
# Select the file from File Dialogue Box
file_name <- file.choose()

# Convert '?' into NA
data <- read.csv(file_name, na.strings = '?')

# Remove the missing values
data<- na.omit(data)

# Removing the column 'id' as it is not required
data<-data[-1]

data$diagnosis <-factor(data$diagnosis)
is.factor(data$diagnosis)
```

```
## [1] TRUE
```

```
## View the dataset
  # View(data)

# Splitting the dataset into training and test data
idx<-sort(sample(nrow(data),as.integer(.70*nrow(data))))

training<-data[idx,]
test<-data[-idx,]

library(e1071)

## Implementing SVM methodology
  # ?svm
  svm.model <- svm(diagnosis ~., data = training)
  svm.pred <- predict(svm.model, test)

  # Frequency table for predictions
  table(actual=test$diagnosis, svm.pred)
```

```
##        svm.pred
## actual   B   M
##      B 101   1
##      M   2  67
```

```
## Calculating the error rate
  # Number of wrong predictions
  SVM_wrong<- (test$diagnosis!=svm.pred)
  # Error Rate in prediction of SVM
  rate<-sum(SVM_wrong)/length(SVM_wrong)

print(paste("Error Rate:" , rate))
```

```
## [1] "Error Rate: 0.0175438596491228"
```