SVM Classification

Mouad & Monish

2022-08-04

```
library(hdrm)
```

```
## Loading required package: glmnet
```

```
## Warning: package 'glmnet' was built under R version 4.2.
```

```
## Loading required package: Matrix
```

```
## Loaded glmnet 4.1-4
```

```
## Loading required package: ncvreg
```

```
## Warning: package 'ncvreg' was built under R version 4.2.
```

```
library(e1071)
```

```
## Warning: package 'e1071' was built under R version 4.2.1
```

```
downloadData(Golub1999)
attachData(Golub1999)
```

About the data set

Before we dig into numbers let's first understand the data

To better understand the data set we generated a handful of

For instance, we have used the cor() function that measures

```
set.seed(1)

n = nrow(X)
# Split data into train and test sets
train_rows <- sample(1:n, n/2)
X.train <- X[train_rows, ]
X.test <- X[-train_rows, ]

y.train <- y[train_rows]
y.test <- y[-train_rows]

y.train = as.numeric(y.train) - 1
y.test = as.numeric(y.test)-1
```

SVM

SVM method has so much to offer in classifying and modelir

1) High performance with a good margin of separation betw
2) Effective in a high dimensional data set
3) Suitable in cases where the number of dimensions is gr
4) Memory efficient

In spite of its powerful capabilities, this tool does have

1) not advisable for large data sets
2) Vulnerable to noises, which may lead target classes to
3) An overabundance of features (predictors) leads to poor

With these characteristics, the SMV method has so many ini

1) data: Golub1999
2) kernel: Used in training and predicting(linear, polynor
3) gamma: Parameter needed for all kernels except linear
4) cost: Cost of constraints violation "C" constant of th

These parameters would be enough to examine and experience