

Data 609 Module 7 HW

Bryan Persaud

5/16/2021

Ex.1

Use the `svm()` algorithm of the `e1071` package to carry out the support vector machine for the `PlantGrowth` data set. Then, discuss the number of support vectors/samples. [Install the `e1071` package in R if needed.]

Solution:

```
library(e1071)
```

```
## Warning: package 'e1071' was built under R version 4.0.5
```

```
plant_svm <- svm(group ~ ., data = PlantGrowth)
summary(plant_svm)
```

```
##
## Call:
## svm(formula = group ~ ., data = PlantGrowth)
##
##
## Parameters:
##   SVM-Type:  C-classification
##   SVM-Kernel: radial
##           cost:  1
##
## Number of Support Vectors:  29
##
##   ( 10 9 10 )
##
##
## Number of Classes:  3
##
## Levels:
##   ctrl trt1 trt2
```

When using the `svm()` algorithm on the `PlantGrowth` data set it shows that there are 29 support vectors.

Ex.2

Do a similar SVM analysis as that in the previous question using the iris data set. Discuss the number of support vectors/samples.

Solution:

```
iris_svm <- svm(Species ~ ., data = iris)
summary(iris_svm)
```

```
##
## Call:
## svm(formula = Species ~ ., data = iris)
##
##
## Parameters:
##   SVM-Type:  C-classification
##   SVM-Kernel: radial
##         cost:  1
##
## Number of Support Vectors:  51
##
##   ( 8 22 21 )
##
##
## Number of Classes:  3
##
## Levels:
##   setosa versicolor virginica
```

When using the svm() algorithm on the iris data set it shows that there are 51 support vectors.

Ex.3

Use the iris data set (or any other data set) to select 80% of the samples for the training svm(), then use the rest 20% for validation. Discuss your results.

Solution:

```
library(caret)
```

```
## Loading required package: lattice
```

```
## Loading required package: ggplot2
```

```
train <- createDataPartition(iris$Species, p = 0.80, list = FALSE)
validation <- iris[-train]
iris_dataset <- data.frame(iris[train])
svm(train ~ ., data = iris_dataset)
```

```
##
## Call:
## svm(formula = train ~ ., data = iris_dataset)
##
##
## Parameters:
##   SVM-Type:  eps-regression
## SVM-Kernel:  radial
##       cost:  1
##       gamma: 0.02857143
##   epsilon:  0.1
##
##
## Number of Support Vectors: 120
```

Selecting 80% of the samples to train `svm()` gives us 120 support vectors.