classification.R

Alexandros

2022-10-16

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.1.2
```

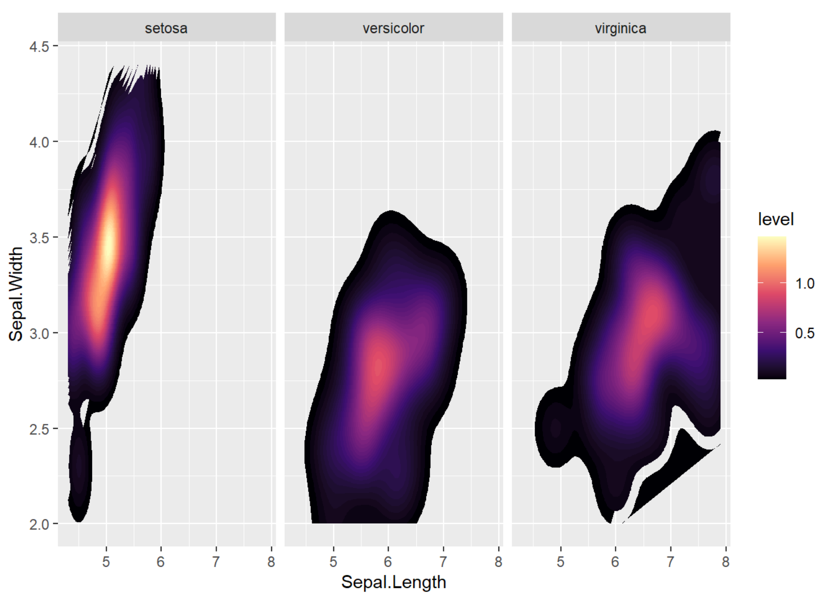
```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.6     v purrr   0.3.4
## v tibble  3.1.2     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1
```

```
## Warning: package 'ggplot2' was built under R version 4.1.3
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(dplyr)
library(ggplot2)



df=iris

###more on ggplot stat https://yjunechoe.github.io/posts/2020-09-26-demystifying-stat-layers-ggplot2/
#stat= 'identity' overides default y and you have to specify y value
##3d and 2d density visualization

## fill=..level.. or stat(nlevel) estimates density kernel



ggplot(df) +
  stat_density_2d(aes(x = Sepal.Length, y = Sepal.Width, fill = ..level..),
                  geom = "polygon", bins = 50,contour = TRUE)+
  facet_wrap(~Species)+    ### multiplot bases on Species
  scale_fill_viridis_c(option = "A")
```



```
### 3d representation with rayshader


#first name ggplot

ggdens=ggplot(df %>% filter(Species=="setosa")) +
  stat_density_2d(aes(x = Sepal.Length, y = Sepal.Width, fill = ..level..),
                  geom = "polygon", bins = 50,contour = TRUE)+
  scale_fill_viridis_c(option = "A")

library(rayshader)
```

```
## Warning: package 'rayshader' was built under R version 4.1.3
```

```
?plot_gg
```

```
## starting httpd help server ...
```

```
##  done
```

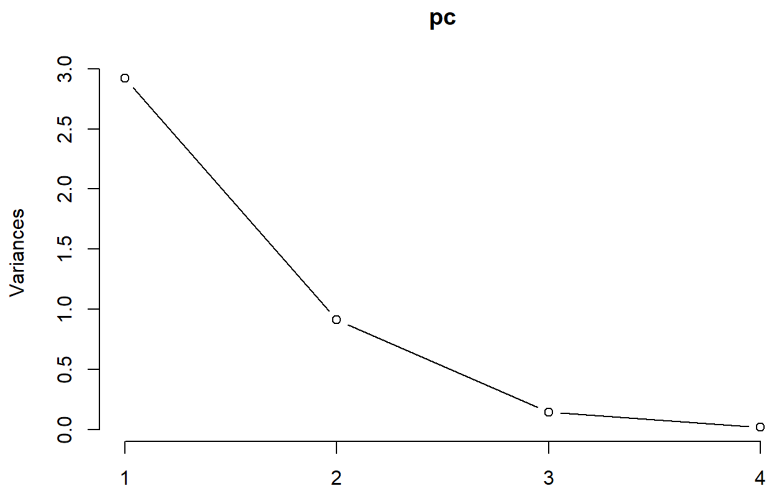
```
#plot_gg(ggdens, width = 5, height = 5, scale = 250,
#        zoom = 0.7, theta = 10, phi = 30, windowsize = c(800, 800))
#Sys.sleep(0.2)
#render_snapshot(clear = TRUE) #captures current rgl view




# Principal components ---------------------------------------------------


pcadata=df[,-5]

# principal comps with data scaling and centered

pc=prcomp(pcadata,
          center=TRUE,
          scale.=TRUE)

attributes(pc)
```

```
## $names
## [1] "sdev"     "rotation" "center"   "scale"    "x"
##
## $class
## [1] "prcomp"
```

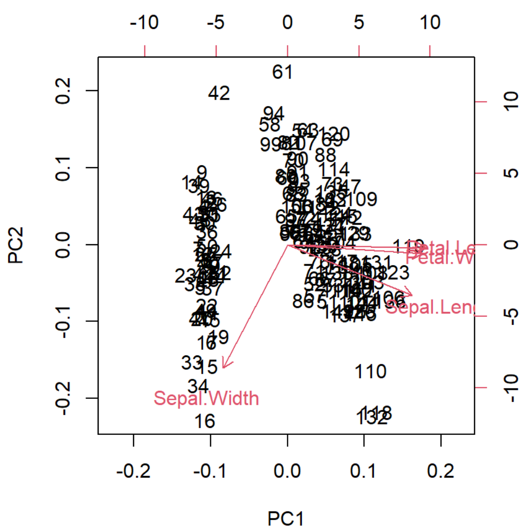
```
print(pc)
```

```
## Standard deviations (1, .., p=4):
## [1] 1.7083611 0.9560494 0.3830886 0.1439265
##
## Rotation (n x k) = (4 x 4):
##                     PC1         PC2        PC3        PC4
## Sepal.Length  0.5210659 -0.37741762  0.7195664  0.2612863
## Sepal.Width  -0.2693474 -0.92329566 -0.2443818 -0.1235096
## Petal.Length  0.5804131 -0.02449161 -0.1421264 -0.8014492
## Petal.Width   0.5648565 -0.06694199 -0.6342727  0.5235971
```

```
screeplot(pc,type="lines") ### See explained variance of PCs
```



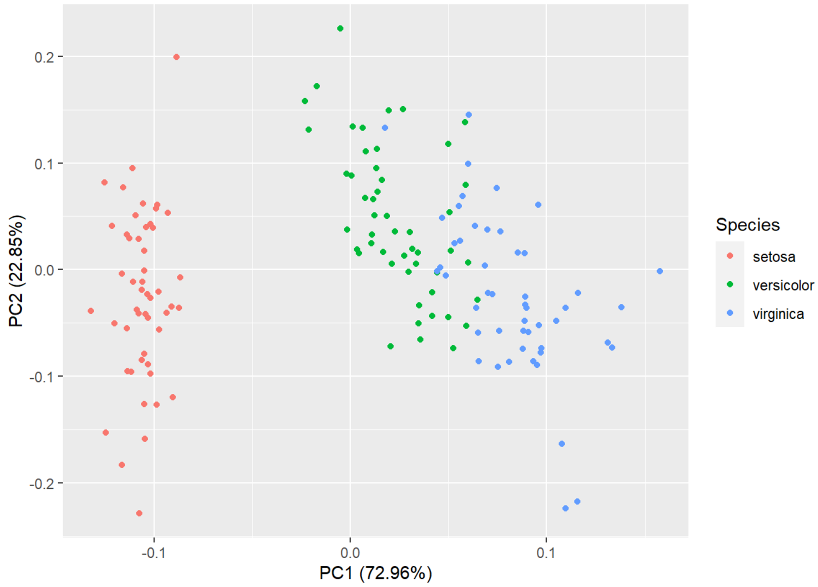

```
biplot(pc) #x axis is first principal component, y axis is second pc
```



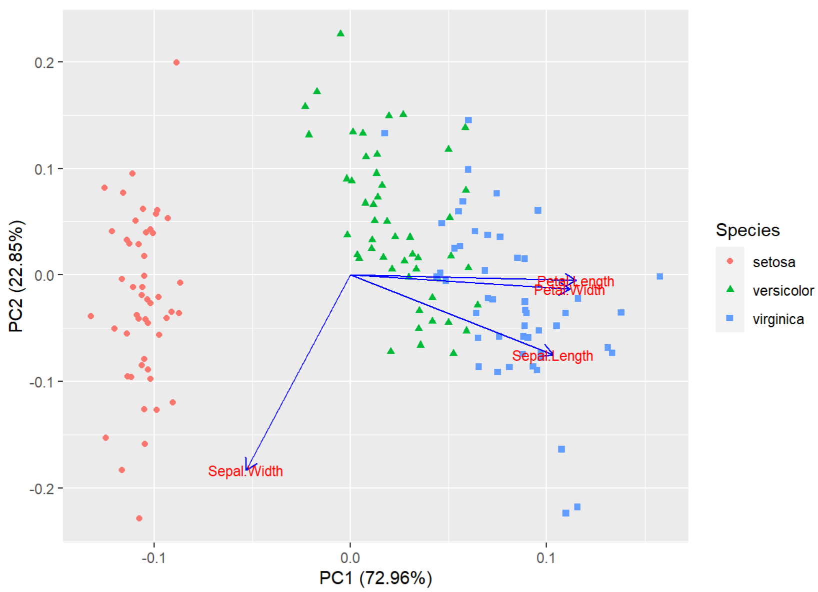

```
### See how pca classifies data
library(ggfortify)
```

```
## Warning: package 'ggfortify' was built under R version 4.1.2
```

```
autoplot(pc,data=df,colour="Species")
```




```r
autoplot(pc, data = df, colour = 'Species',shape="Species",
         loadings = TRUE, loadings.colour = 'blue',
         loadings.label = TRUE, loadings.label.size = 3)
```




```r
# Classification ----------------------------------------------------------

######Data Preperation
library(caret)
```

```
## Warning: package 'caret' was built under R version 4.1.1
```

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

```
##
## Attaching package: 'caret'
```

```
## The following object is masked from 'package:purrr':
##
##     lift
```

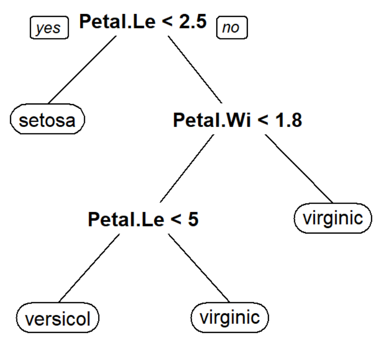
```r
# create training and testing data
parts=createDataPartition(y=df$Species,p=0.8,list=FALSE)
set.seed(66666)
train=df[parts,]
test=df[-parts,]

# specifying the CV technique which will be passed into the train()
#function later and number parameter is the "k" in K-fold

train_control = trainControl(method = "cv", number = 5, search = "grid")


# Decision Tree -----------------------------------------------------------

library(rpart)
```

```
## Warning: package 'rpart' was built under R version 4.1.2
```

```r
library(rpart.plot)
```

```
## Warning: package 'rpart.plot' was built under R version 4.1.3
```

```r
##Decision on full data
# Set minsplit = 2 to fit every data point
full_fit <- rpart(Species ~ ., data = df, minsplit = 2)
prp(full_fit)
```




```r
full_fit$variable.importance
```

```
##  Petal.Width Petal.Length Sepal.Length  Sepal.Width
##     88.96940     85.79403     54.09606     36.01309
```

```r
##Grid search

## create grid for parameters


modelLookup("rpart") #look at parameters of model
```

```
##   model parameter                label forReg forClass probModel
## 1 rpart        cp Complexity Parameter   TRUE     TRUE      TRUE
```

```r
seq(from=0.1,to=1,by=0.1)
```

```
##  [1] 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
```

```r
gs=data.frame(cp=seq(from=0.1,to=1,by=0.1))
gs
```

```
##     cp
## 1  0.1
## 2  0.2
## 3  0.3
## 4  0.4
## 5  0.5
## 6  0.6
## 7  0.7
## 8  0.8
## 9  0.9
## 10 1.0
```

```
iris.tree = train(Species ~ .,
                  data=train,
                  method="rpart",
                  trControl = trainControl(method = "repeatedcv"),
                  tuneGrid=gs,
                  metric="Accuracy")



iris.tree
```

```
## CART
##
## 120 samples
##   4 predictor
##   3 classes: 'setosa', 'versicolor', 'virginica'
##
## No pre-processing
## Resampling: Cross-Validated (10 fold, repeated 1 times)
## Summary of sample sizes: 108, 108, 108, 108, 108, 108, ...
## Resampling results across tuning parameters:
##
##   cp   Accuracy   Kappa
##   0.1  0.8916667  0.8375
##   0.2  0.8916667  0.8375
##   0.3  0.8916667  0.8375
##   0.4  0.8916667  0.8375
##   0.5  0.3333333  0.0000
##   0.6  0.3333333  0.0000
##   0.7  0.3333333  0.0000
##   0.8  0.3333333  0.0000
##   0.9  0.3333333  0.0000
##   1.0  0.3333333  0.0000
##
## Accuracy was used to select the optimal model using the largest value.
## The final value used for the model was cp = 0.4.
```

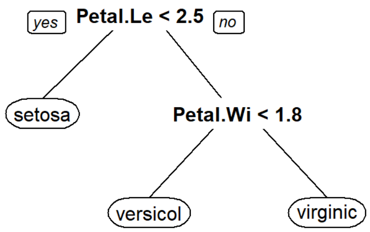
```
iris.tree$bestTune # best tune
```

```
##    cp
## 4 0.4
```

```
iris.tree$finalModel #best model
```

```
## n= 120
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
## 1) root 120 80 setosa (0.3333333 0.3333333 0.3333333)
##   2) Petal.Length< 2.45 40  0 setosa (1.0000000 0.0000000 0.0000000) *
##   3) Petal.Length>=2.45 80 40 versicolor (0.0000000 0.5000000 0.5000000)
##     6) Petal.Width< 1.75 45  5 versicolor (0.0000000 0.8888889 0.1111111) *
##     7) Petal.Width>=1.75 35  0 virginica (0.0000000 0.0000000 1.0000000) *
```

```
prp(iris.tree$finalModel)
```




```
## predict
iris.pred = predict(iris.tree, newdata = test)

table(iris.pred, test$Species) ## table for test data and trained
```

```
##
## iris.pred    setosa versicolor virginica
##   setosa         10          0         0
##   versicolor      0          9         0
##   virginica       0          1        10
```

```
sum(iris.pred==test$Species) #### Count correct classifications
```

```
## [1] 29
```

```
# SVM ---------------------------------------------------------------------


## svm needs scaling

strain=train %>% mutate_at(.vars=c(1:4),scale)
strain
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width    Species
## 1    -0.84999182  1.06464034  -1.31937075 -1.28864901     setosa
## 2    -1.08692334 -0.08425211  -1.31937075 -1.28864901     setosa
## 3    -1.32385485  0.37530487  -1.37565389 -1.28864901     setosa
## 4    -1.44232061  0.14552638  -1.26308760 -1.28864901     setosa
## 5    -0.96845758  1.29441882  -1.31937075 -1.28864901     setosa
## 6    -0.49459454  1.98375429  -1.15052131 -1.02809656     setosa
## 7    -1.44232061  0.83486185  -1.31937075 -1.15837279     setosa
## 8    -0.96845758  0.83486185  -1.26308760 -1.28864901     setosa
## 9    -1.67925213 -0.31403060  -1.31937075 -1.28864901     setosa
## 10   -1.08692334  0.14552638  -1.26308760 -1.41892524     setosa
## 11   -0.49459454  1.52419731  -1.26308760 -1.28864901     setosa
## 12   -1.20538909  0.83486185  -1.20680445 -1.28864901     setosa
## 14   -1.79771789 -0.08425211  -1.48822018 -1.41892524     setosa
## 15   -0.02073151  2.21353278  -1.43193704 -1.28864901     setosa
## 17   -0.49459454  1.98375429  -1.37565389 -1.02809656     setosa
## 20   -0.84999182  1.75397580  -1.26308760 -1.15837279     setosa
## 22   -0.84999182  1.52419731  -1.26308760 -1.02809656     setosa
## 24   -0.84999182  0.60508336  -1.15052131 -0.89782033     setosa
## 25   -1.20538909  0.83486185  -1.03795502 -1.28864901     setosa
## 26   -0.96845758 -0.08425211  -1.20680445 -1.28864901     setosa
## 27   -0.96845758  0.83486185  -1.20680445 -1.02809656     setosa
## 29   -0.73152606  0.83486185  -1.31937075 -1.28864901     setosa
## 30   -1.32385485  0.37530487  -1.20680445 -1.28864901     setosa
## 31   -1.20538909  0.14552638  -1.20680445 -1.28864901     setosa
## 32   -0.49459454  0.83486185  -1.26308760 -1.02809656     setosa
## 33   -0.73152606  2.44331127  -1.26308760 -1.41892524     setosa
## 34   -0.37612878  2.67308976  -1.31937075 -1.28864901     setosa
## 35   -1.08692334  0.14552638  -1.26308760 -1.28864901     setosa
## 36   -0.96845758  0.37530487  -1.43193704 -1.28864901     setosa
## 37   -0.37612878  1.06464034  -1.37565389 -1.28864901     setosa
## 38   -1.08692334  1.29441882  -1.31937075 -1.41892524     setosa
## 39   -1.67925213 -0.08425211  -1.37565389 -1.28864901     setosa
## 41   -0.96845758  1.06464034  -1.37565389 -1.15837279     setosa
## 42   -1.56078637 -1.69270154  -1.37565389 -1.15837279     setosa
## 43   -1.67925213  0.37530487  -1.37565389 -1.28864901     setosa
## 46   -1.20538909 -0.08425211  -1.31937075 -1.15837279     setosa
## 47   -0.84999182  1.75397580  -1.20680445 -1.28864901     setosa
## 48   -1.44232061  0.37530487  -1.31937075 -1.28864901     setosa
## 49   -0.61306030  1.52419731  -1.26308760 -1.28864901     setosa
## 50   -0.96845758  0.60508336  -1.31937075 -1.28864901     setosa
## 51    1.40085760  0.37530487   0.53797307  0.27466571 versicolor
## 53    1.28239184  0.14552638   0.65053936  0.40494194 versicolor
## 54   -0.37612878 -1.69270154   0.14399105  0.14438949 versicolor
## 55    0.80852880 -0.54380909   0.48168992  0.40494194 versicolor
## 56   -0.13919727 -0.54380909   0.42540678  0.14438949 versicolor
## 57    0.57159729  0.60508336   0.53797307  0.53521817 versicolor
## 58   -1.08692334 -1.46292305  -0.24999097 -0.24643920 versicolor
## 60   -0.73152606 -0.77358758   0.08770790  0.27466571 versicolor
## 61   -0.96845758 -2.38203701  -0.13742468 -0.24643920 versicolor
## 62    0.09773425 -0.08425211   0.25655734  0.40494194 versicolor
## 63    0.21620001 -1.92248003   0.14399105 -0.24643920 versicolor
## 64    0.33466577 -0.31403060   0.53797307  0.27466571 versicolor
## 65   -0.25766303 -0.31403060  -0.08114154  0.14438949 versicolor
## 66    1.04546032  0.14552638   0.36912363  0.27466571 versicolor
## 67   -0.25766303 -0.08425211   0.42540678  0.40494194 versicolor
## 68   -0.02073151 -0.77358758   0.20027419 -0.24643920 versicolor
## 69    0.45313153 -1.92248003   0.42540678  0.40494194 versicolor
## 70   -0.25766303 -1.23314456   0.08770790 -0.11616297 versicolor
## 72    0.33466577 -0.54380909   0.14399105  0.14438949 versicolor
## 73    0.57159729 -1.23314456   0.65053936  0.40494194 versicolor
## 74    0.33466577 -0.54380909   0.53797307  0.01411326 versicolor
## 75    0.69006304 -0.31403060   0.31284049  0.14438949 versicolor
## 77    1.16392608 -0.54380909   0.59425622  0.27466571 versicolor
## 78    1.04546032 -0.08425211   0.70682251  0.66549439 versicolor
## 80   -0.13919727 -1.00336607  -0.13742468 -0.24643920 versicolor
## 82   -0.37612878 -1.46292305  -0.02485839 -0.24643920 versicolor
## 85   -0.49459454 -0.08425211   0.42540678  0.40494194 versicolor
## 87    1.04546032  0.14552638   0.53797307  0.40494194 versicolor
## 88    0.57159729 -1.69270154   0.36912363  0.14438949 versicolor
## 89   -0.25766303 -0.08425211   0.20027419  0.14438949 versicolor
## 90   -0.37612878 -1.23314456   0.14399105  0.14438949 versicolor
## 92    0.33466577 -0.08425211   0.48168992  0.27466571 versicolor
## 93   -0.02073151 -1.00336607   0.14399105  0.01411326 versicolor
## 94   -0.96845758 -1.69270154  -0.24999097 -0.24643920 versicolor
## 95   -0.25766303 -0.77358758   0.25655734  0.14438949 versicolor
## 96   -0.13919727 -0.08425211   0.25655734  0.01411326 versicolor
## 97   -0.13919727 -0.31403060   0.25655734  0.14438949 versicolor
## 98    0.45313153 -0.31403060   0.31284049  0.14438949 versicolor
## 99   -0.84999182 -1.23314456  -0.41884041 -0.11616297 versicolor
## 100  -0.13919727 -0.54380909   0.20027419  0.14438949 versicolor
## 101   0.57159729  0.60508336   1.26965397  1.70770421  virginica
## 103   1.51932335 -0.08425211   1.21337082  1.18659930  virginica
## 104   0.57159729 -0.31403060   1.04452138  0.79577062  virginica
```

```
## 105   0.80852880 -0.08425211   1.15708767  1.31687553  virginica
## 106   2.11165215 -0.08425211   1.60735284  1.18659930  virginica
## 107  -1.08692334 -1.23314456   0.42540678  0.66549439  virginica
## 109   1.04546032 -1.23314456   1.15708767  0.79577062  virginica
## 110   1.63778911  1.29441882   1.32593711  1.70770421  virginica
## 111   0.80852880  0.37530487   0.76310565  1.05632307  virginica
## 112   0.69006304 -0.77358758   0.87567194  0.92604685  virginica
## 113   1.16392608 -0.08425211   0.98823824  1.18659930  virginica
## 114  -0.13919727 -1.23314456   0.70682251  1.05632307  virginica
## 115  -0.02073151 -0.54380909   0.76310565  1.57742798  virginica
## 116   0.69006304  0.37530487   0.87567194  1.44715176  virginica
## 117   0.80852880 -0.08425211   0.98823824  0.79577062  virginica
## 118   2.23011791  1.75397580   1.66363599  1.31687553  virginica
## 119   2.23011791 -1.00336607   1.77620228  1.44715176  virginica
## 120   0.21620001 -1.92248003   0.70682251  0.40494194  virginica
## 122  -0.25766303 -0.54380909   0.65053936  1.05632307  virginica
## 123   2.23011791 -0.54380909   1.66363599  1.05632307  virginica
## 124   0.57159729 -0.77358758   0.65053936  0.79577062  virginica
## 125   1.04546032  0.60508336   1.10080453  1.18659930  virginica
## 127   0.45313153 -0.54380909   0.59425622  0.79577062  virginica
## 129   0.69006304 -0.54380909   1.04452138  1.18659930  virginica
## 130   1.63778911 -0.08425211   1.15708767  0.53521817  virginica
## 131   1.87472063 -0.54380909   1.32593711  0.92604685  virginica
## 132   2.46704942  1.75397580   1.49478655  1.05632307  virginica
## 134   0.57159729 -0.54380909   0.76310565  0.40494194  virginica
## 135   0.33466577 -1.00336607   1.04452138  0.27466571  virginica
## 137   0.57159729  0.83486185   1.04452138  1.57742798  virginica
## 138   0.69006304  0.14552638   0.98823824  0.79577062  virginica
## 140   1.28239184  0.14552638   0.93195509  1.18659930  virginica
## 141   1.04546032  0.14552638   1.04452138  1.57742798  virginica
## 142   1.28239184  0.14552638   0.76310565  1.44715176  virginica
## 143  -0.02073151 -0.77358758   0.76310565  0.92604685  virginica
## 144   1.16392608  0.37530487   1.21337082  1.44715176  virginica
## 145   1.04546032  0.60508336   1.10080453  1.70770421  virginica
## 148   0.80852880 -0.08425211   0.81938880  1.05632307  virginica
## 149   0.45313153  0.83486185   0.93195509  1.44715176  virginica
## 150   0.09773425 -0.08425211   0.76310565  0.79577062  virginica
```

```
stest=test %>% mutate_at(.vars=c(1:4),scale)









classifier = train(form = Species ~ ., data = strain, method = 'svmRadial')
classifier
```

```
## Support Vector Machines with Radial Basis Function Kernel
##
## 120 samples
##   4 predictor
##   3 classes: 'setosa', 'versicolor', 'virginica'
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 120, 120, 120, 120, 120, 120, ...
## Resampling results across tuning parameters:
##
##   C     Accuracy   Kappa
##   0.25  0.9477298  0.9206564
##   0.50  0.9543123  0.9304854
##   1.00  0.9537345  0.9295779
##
## Tuning parameter 'sigma' was held constant at a value of 0.4707987
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were sigma = 0.4707987 and C = 0.5.
```

```
classifier$bestTune
```

```
##       sigma   C
## 2 0.4707987 0.5
```

```
##create a grid with map functions

modelLookup("svmRadial")
```

```
##       model parameter label forReg forClass probModel
## 1 svmRadial     sigma Sigma   TRUE     TRUE      TRUE
## 2 svmRadial         C  Cost   TRUE     TRUE      TRUE
```

```
sigmas=seq(0.1,1,by=0.2)
cs=seq(0.1,1,by=0.2)
gs=expand.grid(sigmas,cs)
gs
```

```
##    Var1 Var2
## 1   0.1  0.1
## 2   0.3  0.1
## 3   0.5  0.1
## 4   0.7  0.1
## 5   0.9  0.1
## 6   0.1  0.3
## 7   0.3  0.3
## 8   0.5  0.3
## 9   0.7  0.3
## 10  0.9  0.3
## 11  0.1  0.5
## 12  0.3  0.5
## 13  0.5  0.5
## 14  0.7  0.5
## 15  0.9  0.5
## 16  0.1  0.7
## 17  0.3  0.7
## 18  0.5  0.7
## 19  0.7  0.7
## 20  0.9  0.7
## 21  0.1  0.9
## 22  0.3  0.9
## 23  0.5  0.9
## 24  0.7  0.9
## 25  0.9  0.9
```

```
names(gs)=c("sigma","C")


##hyper parameter tuning
classifier = train(form = Species ~ ., data = strain, method = 'svmRadial',
                   tuneGrid=gs )


classifier
```

```
## Support Vector Machines with Radial Basis Function Kernel
##
## 120 samples
##   4 predictor
##   3 classes: 'setosa', 'versicolor', 'virginica'
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 120, 120, 120, 120, 120, 120, ...
## Resampling results across tuning parameters:
##
##   sigma  C    Accuracy   Kappa
##   0.1    0.1  0.8521709  0.7830846
##   0.1    0.3  0.9183085  0.8769578
##   0.1    0.5  0.9414813  0.9116723
##   0.1    0.7  0.9484058  0.9219536
##   0.1    0.9  0.9581111  0.9366987
##   0.3    0.1  0.8766271  0.8160686
##   0.3    0.3  0.9377889  0.9059394
##   0.3    0.5  0.9555595  0.9326931
##   0.3    0.7  0.9575680  0.9356785
##   0.3    0.9  0.9602447  0.9397637
##   0.5    0.1  0.8892201  0.8352790
##   0.5    0.3  0.9469689  0.9197462
##   0.5    0.5  0.9522649  0.9277589
##   0.5    0.7  0.9548814  0.9317020
##   0.5    0.9  0.9602250  0.9398056
##   0.7    0.1  0.8881251  0.8339909
##   0.7    0.3  0.9437214  0.9147331
##   0.7    0.5  0.9478417  0.9210599
##   0.7    0.7  0.9523737  0.9278781
##   0.7    0.9  0.9549861  0.9318585
##   0.9    0.1  0.8794258  0.8214894
##   0.9    0.3  0.9417253  0.9117572
##   0.9    0.5  0.9479269  0.9211816
##   0.9    0.7  0.9487603  0.9224314
##   0.9    0.9  0.9527087  0.9284494
##
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were sigma = 0.3 and C = 0.9.
```

```
classifier$bestTune
```

```
##    sigma   C
## 10   0.3 0.9
```

```
###hyper parameter tuning AND cross validation
trControl = trainControl(method = "repeatedcv")
classifier = train(form = Species ~ ., data = strain, method = 'svmRadial',
                   tuneGrid=gs,
                   trControl = trainControl(method = "cv"))


classifier$bestTune
```

```
##   sigma   C
## 9   0.3 0.7
```

```
#predicting resuklts
y_pred = predict(classifier, newdata = stest)

y_pred
```

```
##  [1] setosa     setosa     setosa     setosa     setosa     setosa    
##  [7] setosa     setosa     setosa     setosa     versicolor versicolor
## [13] versicolor versicolor versicolor versicolor versicolor virginica 
## [19] versicolor versicolor virginica  virginica  virginica  virginica 
## [25] virginica  virginica  virginica  virginica  virginica  virginica 
## Levels: setosa versicolor virginica
```

```
#confusion matrix for train results
modelsvm=classifier$finalModel

modelsvm@fitted
```

```
##   [1] setosa     setosa     setosa     setosa     setosa     setosa    
##   [7] setosa     setosa     setosa     setosa     setosa     setosa    
##  [13] setosa     setosa     setosa     setosa     setosa     setosa    
##  [19] setosa     setosa     setosa     setosa     setosa     setosa    
##  [25] setosa     setosa     setosa     setosa     setosa     setosa    
##  [31] setosa     setosa     setosa     setosa     setosa     setosa    
##  [37] setosa     setosa     setosa     setosa     versicolor versicolor
##  [43] versicolor versicolor versicolor versicolor versicolor versicolor
##  [49] versicolor versicolor versicolor versicolor versicolor versicolor
##  [55] versicolor versicolor versicolor versicolor versicolor versicolor
##  [61] versicolor versicolor versicolor virginica  versicolor versicolor
##  [67] versicolor versicolor versicolor versicolor versicolor versicolor
##  [73] versicolor versicolor versicolor versicolor versicolor versicolor
##  [79] versicolor versicolor virginica  virginica  virginica  virginica 
##  [85] virginica  virginica  virginica  virginica  virginica  virginica 
##  [91] virginica  virginica  virginica  virginica  virginica  virginica 
##  [97] virginica  versicolor virginica  virginica  virginica  virginica 
## [103] virginica  virginica  virginica  virginica  virginica  versicolor
## [109] virginica  virginica  virginica  virginica  virginica  virginica 
## [115] virginica  virginica  virginica  virginica  virginica  virginica 
## Levels: setosa versicolor virginica
```

```
table(strain$Species, modelsvm@fitted)
```

```
## 
##              setosa versicolor virginica
##   setosa         40          0         0
##   versicolor      0         39         1
##   virginica       0          2        38
```

```
#confussion matrix for test results

classifier$results
```

```
##    sigma   C  Accuracy  Kappa AccuracySD    KappaSD
## 1    0.1 0.1 0.8750000 0.8125 0.05892557 0.08838835
## 2    0.1 0.3 0.9333333 0.9000 0.05270463 0.07905694
## 3    0.1 0.5 0.9500000 0.9250 0.04303315 0.06454972
## 4    0.1 0.7 0.9583333 0.9375 0.04392052 0.06588078
## 5    0.1 0.9 0.9666667 0.9500 0.04303315 0.06454972
## 6    0.3 0.1 0.9000000 0.8500 0.07657805 0.11486707
## 7    0.3 0.3 0.9500000 0.9250 0.05826716 0.08740074
## 8    0.3 0.5 0.9583333 0.9375 0.05892557 0.08838835
## 9    0.3 0.7 0.9666667 0.9500 0.04303315 0.06454972
## 10   0.3 0.9 0.9666667 0.9500 0.04303315 0.06454972
## 11   0.5 0.1 0.9166667 0.8750 0.06804138 0.10206207
## 12   0.5 0.3 0.9500000 0.9250 0.08050765 0.12076147
## 13   0.5 0.5 0.9500000 0.9250 0.08050765 0.12076147
## 14   0.5 0.7 0.9583333 0.9375 0.08098544 0.12147816
## 15   0.5 0.9 0.9583333 0.9375 0.05892557 0.08838835
## 16   0.7 0.1 0.9250000 0.8875 0.06148873 0.09223310
## 17   0.7 0.3 0.9500000 0.9250 0.08050765 0.12076147
## 18   0.7 0.5 0.9500000 0.9250 0.08050765 0.12076147
## 19   0.7 0.7 0.9500000 0.9250 0.08050765 0.12076147
## 20   0.7 0.9 0.9583333 0.9375 0.08098544 0.12147816
## 21   0.9 0.1 0.9333333 0.9000 0.06573422 0.09860133
## 22   0.9 0.3 0.9416667 0.9125 0.07905694 0.11858541
## 23   0.9 0.5 0.9416667 0.9125 0.07905694 0.11858541
## 24   0.9 0.7 0.9583333 0.9375 0.08098544 0.12147816
## 25   0.9 0.9 0.9500000 0.9250 0.08050765 0.12076147
```

```
cm = table(stest$Species, y_pred)
cm
```

```
##             y_pred
##              setosa versicolor virginica
##   setosa         10          0         0
##   versicolor      0          9         1
##   virginica       0          0        10
```

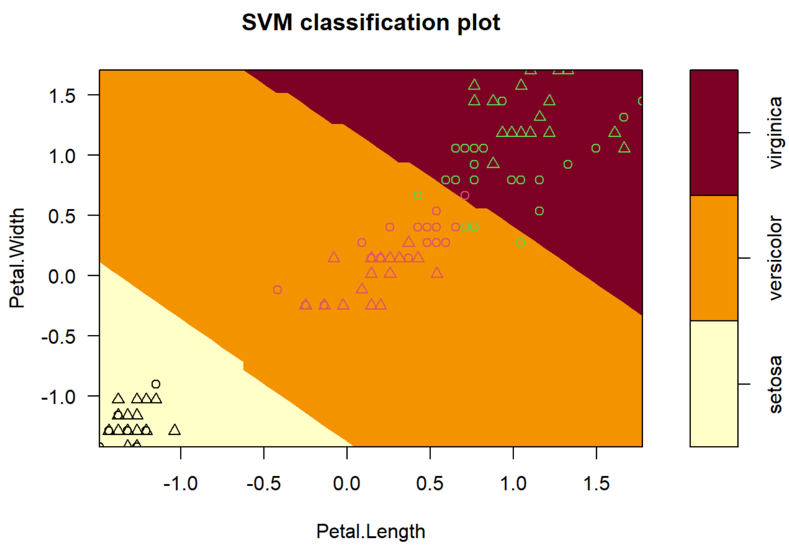
```
##visualize training results
library(e1071)
```

```
## Warning: package 'e1071' was built under R version 4.1.2
```

```
classifier$bestTune
```

```
##   sigma   C
## 9   0.3 0.7
```

```
svmfit <- svm(Species~., data = strain, kernel = "radial",sigma=0.1,cost=0.7)



plot(svmfit, strain, Petal.Width ~ Petal.Length,svSymbol =1,dataSymbol=2)
```




```
df$Sepal.Length %>% length
```

```
## [1] 150
```

```
# Understaning map functions ----------------------------------------------


df %>% split(.$Species)    #split dataframe into 3 based on species
```

```
## $setosa
##    Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1           5.1         3.5          1.4         0.2  setosa
## 2           4.9         3.0          1.4         0.2  setosa
## 3           4.7         3.2          1.3         0.2  setosa
## 4           4.6         3.1          1.5         0.2  setosa
## 5           5.0         3.6          1.4         0.2  setosa
## 6           5.4         3.9          1.7         0.4  setosa
## 7           4.6         3.4          1.4         0.3  setosa
## 8           5.0         3.4          1.5         0.2  setosa
## 9           4.4         2.9          1.4         0.2  setosa
## 10          4.9         3.1          1.5         0.1  setosa
## 11          5.4         3.7          1.5         0.2  setosa
## 12          4.8         3.4          1.6         0.2  setosa
## 13          4.8         3.0          1.4         0.1  setosa
## 14          4.3         3.0          1.1         0.1  setosa
## 15          5.8         4.0          1.2         0.2  setosa
## 16          5.7         4.4          1.5         0.4  setosa
## 17          5.4         3.9          1.3         0.4  setosa
## 18          5.1         3.5          1.4         0.3  setosa
## 19          5.7         3.8          1.7         0.3  setosa
## 20          5.1         3.8          1.5         0.3  setosa
## 21          5.4         3.4          1.7         0.2  setosa
## 22          5.1         3.7          1.5         0.4  setosa
## 23          4.6         3.6          1.0         0.2  setosa
## 24          5.1         3.3          1.7         0.5  setosa
## 25          4.8         3.4          1.9         0.2  setosa
## 26          5.0         3.0          1.6         0.2  setosa
## 27          5.0         3.4          1.6         0.4  setosa
## 28          5.2         3.5          1.5         0.2  setosa
## 29          5.2         3.4          1.4         0.2  setosa
## 30          4.7         3.2          1.6         0.2  setosa
## 31          4.8         3.1          1.6         0.2  setosa
## 32          5.4         3.4          1.5         0.4  setosa
## 33          5.2         4.1          1.5         0.1  setosa
## 34          5.5         4.2          1.4         0.2  setosa
## 35          4.9         3.1          1.5         0.2  setosa
## 36          5.0         3.2          1.2         0.2  setosa
## 37          5.5         3.5          1.3         0.2  setosa
## 38          4.9         3.6          1.4         0.1  setosa
## 39          4.4         3.0          1.3         0.2  setosa
## 40          5.1         3.4          1.5         0.2  setosa
## 41          5.0         3.5          1.3         0.3  setosa
## 42          4.5         2.3          1.3         0.3  setosa
## 43          4.4         3.2          1.3         0.2  setosa
## 44          5.0         3.5          1.6         0.6  setosa
## 45          5.1         3.8          1.9         0.4  setosa
## 46          4.8         3.0          1.4         0.3  setosa
## 47          5.1         3.8          1.6         0.2  setosa
## 48          4.6         3.2          1.4         0.2  setosa
## 49          5.3         3.7          1.5         0.2  setosa
## 50          5.0         3.3          1.4         0.2  setosa
## 
## $versicolor
##     Sepal.Length Sepal.Width Petal.Length Petal.Width    Species
## 51           7.0         3.2          4.7         1.4 versicolor
## 52           6.4         3.2          4.5         1.5 versicolor
## 53           6.9         3.1          4.9         1.5 versicolor
## 54           5.5         2.3          4.0         1.3 versicolor
## 55           6.5         2.8          4.6         1.5 versicolor
## 56           5.7         2.8          4.5         1.3 versicolor
## 57           6.3         3.3          4.7         1.6 versicolor
## 58           4.9         2.4          3.3         1.0 versicolor
## 59           6.6         2.9          4.6         1.3 versicolor
## 60           5.2         2.7          3.9         1.4 versicolor
## 61           5.0         2.0          3.5         1.0 versicolor
## 62           5.9         3.0          4.2         1.5 versicolor
## 63           6.0         2.2          4.0         1.0 versicolor
## 64           6.1         2.9          4.7         1.4 versicolor
## 65           5.6         2.9          3.6         1.3 versicolor
## 66           6.7         3.1          4.4         1.4 versicolor
## 67           5.6         3.0          4.5         1.5 versicolor
## 68           5.8         2.7          4.1         1.0 versicolor
## 69           6.2         2.2          4.5         1.5 versicolor
## 70           5.6         2.5          3.9         1.1 versicolor
## 71           5.9         3.2          4.8         1.8 versicolor
## 72           6.1         2.8          4.0         1.3 versicolor
## 73           6.3         2.5          4.9         1.5 versicolor
## 74           6.1         2.8          4.7         1.2 versicolor
## 75           6.4         2.9          4.3         1.3 versicolor
## 76           6.6         3.0          4.4         1.4 versicolor
## 77           6.8         2.8          4.8         1.4 versicolor
## 78           6.7         3.0          5.0         1.7 versicolor
## 79           6.0         2.9          4.5         1.5 versicolor
```

```
## 80          5.7         2.6          3.5         1.0 versicolor
## 81          5.5         2.4          3.8         1.1 versicolor
## 82          5.5         2.4          3.7         1.0 versicolor
## 83          5.8         2.7          3.9         1.2 versicolor
## 84          6.0         2.7          5.1         1.6 versicolor
## 85          5.4         3.0          4.5         1.5 versicolor
## 86          6.0         3.4          4.5         1.6 versicolor
## 87          6.7         3.1          4.7         1.5 versicolor
## 88          6.3         2.3          4.4         1.3 versicolor
## 89          5.6         3.0          4.1         1.3 versicolor
## 90          5.5         2.5          4.0         1.3 versicolor
## 91          5.5         2.6          4.4         1.2 versicolor
## 92          6.1         3.0          4.6         1.4 versicolor
## 93          5.8         2.6          4.0         1.2 versicolor
## 94          5.0         2.3          3.3         1.0 versicolor
## 95          5.6         2.7          4.2         1.3 versicolor
## 96          5.7         3.0          4.2         1.2 versicolor
## 97          5.7         2.9          4.2         1.3 versicolor
## 98          6.2         2.9          4.3         1.3 versicolor
## 99          5.1         2.5          3.0         1.1 versicolor
## 100         5.7         2.8          4.1         1.3 versicolor
## 
## $virginica
##     Sepal.Length Sepal.Width Petal.Length Petal.Width   Species
## 101          6.3         3.3          6.0         2.5 virginica
## 102          5.8         2.7          5.1         1.9 virginica
## 103          7.1         3.0          5.9         2.1 virginica
## 104          6.3         2.9          5.6         1.8 virginica
## 105          6.5         3.0          5.8         2.2 virginica
## 106          7.6         3.0          6.6         2.1 virginica
## 107          4.9         2.5          4.5         1.7 virginica
## 108          7.3         2.9          6.3         1.8 virginica
## 109          6.7         2.5          5.8         1.8 virginica
## 110          7.2         3.6          6.1         2.5 virginica
## 111          6.5         3.2          5.1         2.0 virginica
## 112          6.4         2.7          5.3         1.9 virginica
## 113          6.8         3.0          5.5         2.1 virginica
## 114          5.7         2.5          5.0         2.0 virginica
## 115          5.8         2.8          5.1         2.4 virginica
## 116          6.4         3.2          5.3         2.3 virginica
## 117          6.5         3.0          5.5         1.8 virginica
## 118          7.7         3.8          6.7         2.2 virginica
## 119          7.7         2.6          6.9         2.3 virginica
## 120          6.0         2.2          5.0         1.5 virginica
## 121          6.9         3.2          5.7         2.3 virginica
## 122          5.6         2.8          4.9         2.0 virginica
## 123          7.7         2.8          6.7         2.0 virginica
## 124          6.3         2.7          4.9         1.8 virginica
## 125          6.7         3.3          5.7         2.1 virginica
## 126          7.2         3.2          6.0         1.8 virginica
## 127          6.2         2.8          4.8         1.8 virginica
## 128          6.1         3.0          4.9         1.8 virginica
## 129          6.4         2.8          5.6         2.1 virginica
## 130          7.2         3.0          5.8         1.6 virginica
## 131          7.4         2.8          6.1         1.9 virginica
## 132          7.9         3.8          6.4         2.0 virginica
## 133          6.4         2.8          5.6         2.2 virginica
## 134          6.3         2.8          5.1         1.5 virginica
## 135          6.1         2.6          5.6         1.4 virginica
## 136          7.7         3.0          6.1         2.3 virginica
## 137          6.3         3.4          5.6         2.4 virginica
## 138          6.4         3.1          5.5         1.8 virginica
## 139          6.0         3.0          4.8         1.8 virginica
## 140          6.9         3.1          5.4         2.1 virginica
## 141          6.7         3.1          5.6         2.4 virginica
## 142          6.9         3.1          5.1         2.3 virginica
## 143          5.8         2.7          5.1         1.9 virginica
## 144          6.8         3.2          5.9         2.3 virginica
## 145          6.7         3.3          5.7         2.5 virginica
## 146          6.7         3.0          5.2         2.3 virginica
## 147          6.3         2.5          5.0         1.9 virginica
## 148          6.5         3.0          5.2         2.0 virginica
## 149          6.2         3.4          5.4         2.3 virginica
## 150          5.9         3.0          5.1         1.8 virginica
```

```
df %>% split(.$Species) %>%    #split dataframe into 3 based on species
  map(~lm(Petal.Length~Sepal.Length,data=.x)) %>%   # run an lm regression on each split dataframe
  map(summary) %>%  #Get summary for every model
  map("r.squared") #get the R squared of each regression
```

```
## $setosa
## [1] 0.07138289
##
## $versicolor
## [1] 0.5685898
##
## $virginica
## [1] 0.7468844
```

```
#expand grid makes data frame out of combinations

sigmas=seq(0.1,1,by=0.2)
cs=seq(0.1,1,by=0.2)
gs=expand.grid(sigmas,cs)
```