Support Vector Machines and Neural Networks

Purpose of this Project:

This project involves building SVM and neural network regression models to answer a number of questions. We will use the Carseats dataset that is part of the ISLR package.

Loading required libraries:

```
library(ISLR)
library(dplyr)
library(glmnet)
library(caret)
```

Selecting required features:

```
Carseats_Filtered <- Carseats %>% select("Sales", "Price",
"Advertising","Population","Age","Income","Education")
```

Building a linear SVM regression model to predict Sales based on all other attributes ("Price", "Advertising", "Population", "Age", "Income" and "Education").:

```
set.seed(1203)
SVM_Model<- train(Sales~.,data=Carseats_Filtered,method="svmLinear",preProcess=c("center","scale"),tunel
SVM_Model
```

```
## Support Vector Machines with Linear Kernel
##
## 400 samples
##   6 predictor
##
## Pre-processing: centered (6), scaled (6)
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 400, 400, 400, 400, 400, 400, ...
## Resampling results:
##
##   RMSE      Rsquared   MAE
##   2.285293  0.3438381  1.83564
##
## Tuning parameter 'C' was held constant at a value of 1
```

R square is 34.38%

Customizing the search grid by checking the model's performance for C parameter of 0.1,.5,1 and 10 using 2 repeats of 5-fold cross validation.

```r
set.seed(1203)

grid = expand.grid(C= c(0.1,0.5,1,10))

trctrl <- trainControl(method = "repeatedcv", number = 5, repeats = 2)

SVM_Model2 <- train(Sales ~., data = Carseats_Filtered, method = "svmLinear",
trControl=trctrl,
preProcess = c("center", "scale"), tuneGrid = grid,
tuneLength = 10)

SVM_Model2
```

```
## Support Vector Machines with Linear Kernel
##
## 400 samples
##   6 predictor
##
## Pre-processing: centered (6), scaled (6)
## Resampling: Cross-Validated (5 fold, repeated 2 times)
## Summary of sample sizes: 321, 320, 320, 319, 320, 320, ...
## Resampling results across tuning parameters:
##
##   C     RMSE      Rsquared   MAE
##    0.1  2.273056  0.3585139  1.824939
##    0.5  2.270436  0.3598841  1.822315
##    1.0  2.270336  0.3601421  1.822068
##   10.0  2.269440  0.3604274  1.821462
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was C = 10.
```

Training a neural network model to predict Sales based on all other attributes ("Price", "Advertising", "Population", "Age", "Income" and "Education").

```r
Normalization <- preProcess(Carseats_Filtered[,2:7],method = c("center","scale"))
Norm_data<-predict(Normalization,Carseats_Filtered)

NNET_Model<- train(Sales~.,data=Norm_data,method="nnet",linout=TRUE, trace = FALSE)
NNET_Model
```

```
## Neural Network
##
## 400 samples
##   6 predictor
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 400, 400, 400, 400, 400, 400, ...
## Resampling results across tuning parameters:
##
##   size  decay  RMSE      Rsquared   MAE
```

```
##   1     0e+00  2.409339  0.2962855  1.922183
##   1     1e-04  2.445182  0.2794252  1.936612
##   1     1e-01  2.320698  0.3343588  1.848641
##   3     0e+00  2.581837  0.2310638  2.075329
##   3     1e-04  2.802493  0.2345802  2.096510
##   3     1e-01  2.467113  0.2743319  1.970831
##   5     0e+00  2.816794  0.1847461  2.216978
##   5     1e-04  2.698980  0.1957055  2.169599
##   5     1e-01  2.609999  0.2226334  2.108031
##
## RMSE was used to select the optimal model using the smallest value.
## The final values used for the model were size = 1 and decay = 0.1.
```

R-square value with best hyperparameters(size=1) is 33.43.

Consider the following input: Sales=9,Price=6.54,Population=124,Advertising=0,Age=76,Income=110,Education=10. Estimating Sales for this record using the above neuralnet model?

```
Input <- data.frame(Price = 6.54, Population = 124, Advertising = 0, Age = 76, Income = 110, Education =

Prediction<- predict(NNET_Model,Input )
Prediction
```

```
##        1
## 4.976281
```

Estimated sales: 4.976281