Labsheet - 6

SVM with Kernels

Machine Learning

BITS F464

I Semester 2021-22

Non Linear SVM

we use the svm() function, however now we can experiment with non-linear kernels. For polynomial kernels we use the parameter degree to adjust the polynomial order. For radial kernels we use the gamma parameter to adjust the y value.

```
# Load iris dataset
data(iris)
```

The data set has 150 observations and 5 variables. First four variables are numeric while the last variable is factor with 3 levels.

Looking at the summary of the data set.

```
summary(iris)
```

We partition the data set into training and testing after applying PCA.

```
#Apply pca on iris dataset and choose two components
pc <- prcomp(iris[,-5],center = T,scale. = T)
pc1<-data.frame(pc$x[,1])
pc2<-data.frame(pc$x[,2])

#partition the data set into training and testing
ind<-sample(2,nrow(pc1),replace = TRUE,prob = c(0.8,0.2))
traindata1<-pc1[ind==1,]
traindata2<-pc2[ind==1,]
train_class<-data.frame(iris[ind==1,5])
testdata1<-pc1[ind==2,]
testdata2<-pc2[ind==2,]
test_class<-data.frame(iris[ind==2,5])

#Rebuild data in list form
```

```
data<-data.frame(class=unlist(train_class), x=unlist(traindata1), y=un
list(traindata2))
```

Create SVM model1, model2, model3 by using radial, polynomial, and sigmoid kernel functions. And analyse their summary and plot the decision boundaries.

RBF kernel :- The RBF kernel function for two points X_1 and X_2 computes the similarity or how close they are to each other. This kernel can be mathematically represented as follows:

$$K(X_1, X_2) = exp(-\frac{||X_1 - X_2||^2}{2\sigma^2})$$

where,

1. 'σ' is the variance and our hyperparameter

2. $||X_1 - X_2||$ is the Euclidean (L_2-norm) Distance between two points X_1 and X_2

```
model1<-svm(class ~ x+y, data=data,
            kernel="radial")
summary(model1)
plot(model1,data)
```

#polynomial kernel.

The polynomial kernel is a kernel function commonly used with support vector machines (SVMs) and other kernelized models, that represents the similarity of vectors (training samples) in a feature space over polynomials of the original variables, allowing learning of non-linear models**.**

$$K(x, y) = (x^t y + c)^d$$

By using the degree parameter, you can change the degree of the polynomial model.

```
model2<-svm(class ~ x+y, data=data,
            kernel="polynomial")
summary(model2)
plot(model2,data)
```

#sigmoid kernel.

It is mostly preferred for neural networks. This kernel function is similar to a two-layer perceptron model of the neural network, which works as an activation function for neurons.

It can be shown as,

$$K(x, y) = \tanh(\gamma . x^t y + r)$$

```
model3<-svm(class ~ x+y, data=data,
            kernel="sigmoid")
summary(model3)
plot(model3,data)
```

Find the accuracy for the three kernels on the test data. Tabulate your findings. Also, compare the non-linear kernels with linear kernel.

```
data_test<-data.frame(x=unlist(testdata1), y=unlist(testdata2))
pred = predict(model1 ,data_test)
```

Exercise

- Try all kernel function on SVM model and predict whether a given car gets high or low gas mileage based on the Auto dataset. Analyze the difference between them. And find which model get highest accuracy.
- Did you realize that IRIS dataset has 3 classes? SVM is a binary classifier by default. Figure out how SVM could work on 3 classes. In general, try to find out how a binary classifier can be used to solve multi-class classification problems.
