R Commands for Classification 1

PARTITIONING DATASETS

```
# To make sure character variables are factor type data
> iris = mutate_if(iris, is.character, as.factor)
# Setting seed for reproducibility
> set.seed(1234)
# Random splitting of iris data as 70% train and 30%test datasets
> index <- createDataPartition(iris$class, p=0.70, list=FALSE)
> trainData = iris[index,]
> testData = iris[-index,]
```

TRAINING MODELS

```
# Build the model
> model_knn = train(trainData[1:4], trainData$class,
method='knn')
# Train the model with preprocessing
> model_knn_pp <- train(trainData[1:4], trainData$class,
method='knn', preProcess=c("center", "scale"))
# To inspect the model trained, just call it
> model_knn
# Training the model by invoking KNN directly to choose the k
value
> model_knn1 = knn3(trainData[,1:4], testData$class,
cl=trainData$class, k=1)
```

EVALUATING MODELS

```
# Predict the labels of the test set
> predictions<-predict(object=model_k1,testData[,1:4])
# Evaluate the predictions
> table(predictions)
# Confusion matrix
> confusionMatrix(predictions,testData$class)
```

USING K-FOLD CROSS VALIDATION

```
# define training control
train_control <- trainControl(method="cv", number=10)
# fix the parameters of the algorithm
grid <- expand.grid(k = c(5,7,9,15,19,21))
# train the model
model_cv <- train(trainData[1:4], trainData$class,
trControl=train_control, method="knn", tuneGrid=grid)
```