**CSE – 6005 – Machine Learning**

**Lab Experiment – 13 - Comparison of Machine Learning Algorithms.**

# load libraries

library(mlbench)

library(caret)

# load the dataset

data(PimaIndiansDiabetes)

# prepare training scheme

control <- trainControl(method="repeatedcv", number=10, repeats=3)

# CART

set.seed(7)

fit.cart <- train(diabetes~., data=PimaIndiansDiabetes, method="rpart", trControl=control)

# LDA

set.seed(7)

fit.lda <- train(diabetes~., data=PimaIndiansDiabetes, method="lda", trControl=control)

# SVM

set.seed(7)

fit.svm <- train(diabetes~., data=PimaIndiansDiabetes, method="svmRadial", trControl=control)

# kNN

set.seed(7)

fit.knn <- train(diabetes~., data=PimaIndiansDiabetes, method="knn", trControl=control)

# Random Forest

set.seed(7)

fit.rf <- train(diabetes~., data=PimaIndiansDiabetes, method="rf", trControl=control)

# collect resamples

results <- resamples(list(CART=fit.cart, LDA=fit.lda, SVM=fit.svm, KNN=fit.knn, RF=fit.rf))

summary(results)

# box and whisker plots to compare models

scales <- list(x=list(relation="free"), y=list(relation="free"))

bwplot(results, scales=scales)

scales <- list(x=list(relation="free"), y=list(relation="free"))

densityplot(results, scales=scales, pch = "|")

scales <- list(x=list(relation="free"), y=list(relation="free"))

dotplot(results, scales=scales)

parallelplot(results)

splom(results)

xyplot(results, models=c("LDA", "SVM"))

# difference in model predictions

diffs <- diff(results)

# summarize p-values for pair-wise comparisons

summary(diffs)







