DATA 624: Project 2

DATA 624 - Predictive Analytics

Group 2

Group Members:

Vinicio Haro

Sang Yoon (Andy) Hwang

Julian McEachern

Jeremy O'Brien

Bethany Poulin

10 December 2019

Contents

Appendix

```
library(party)
library(partykit)
library(caret)
library(MLmetrics)
#install.packages("rattle")
library(rattle)


# Model Performance
*  Set1 = Caret: bagImputed; no additional pre-processing
*  Set2 = Caret: bagImputed; PreP `method=c('center', 'scale', 'nzv', 'BoxCox')`

#### Train Performance:
tbl.perf.train1 %>%
  kable(caption="Train1 Performance", booktabs=T, digits=4) %>%
  kable_styling()
tbl.perf.train2 %>%
  kable(caption="Train2 Performance", booktabs=T, digits=4) %>%
  kable_styling()

#### Test Accuracy:
tbl.perf.test1 %>%
  kable(caption="Test1 Performance", booktabs=T, digits=4) %>%
  kable_styling()
tbl.perf.test2 %>%
  kable(caption="Test2 Performance", booktabs=T, digits=4) %>%
  kable_styling()

# New Model : RPART

##Metrics
set.seed(58677)
grid_rpart <- expand.grid(maxdepth = 1:20)
fit.rpart <- train(PH~., data=train_trans, metric="RMSE", method = "rpart2", tuneGrid = grid_rpart, tun
rp.PERF <- rbind(getTrainPerf(fit.rpart))
rp.MAPE <- cbind(MAPE = MAPE(fit.rpart$pred$pred, fit.rpart$pred$obs))
rp.ACC <- cbind(rp.MAPE, rp.PERF)
fit.rpart2 <- train(PH~., data=train_trans, metric="RMSE", method = "rpart", tuneLength=tl, trControl=t
rp.PERF2 <- rbind(getTrainPerf(fit.rpart2))
rp.MAPE2 <- cbind(MAPE = MAPE(fit.rpart2$pred$pred, fit.rpart2$pred$obs))
rp.ACC2 <- cbind(rp.MAPE2, rp.PERF2)
rbind(rp.ACC, rp.ACC2) %>% kable() %>% kable_styling()

##Tune Grid
p1<-ggplot(fit.rpart) + theme_bw() + theme() + labs(title = "Cart1 CV TUNE GRID")
p2<-ggplot(fit.rpart2) + theme_bw() + theme() + labs(title = "Cart2 CV TUNE GRID")
```

```
gridExtra::grid.arrange(p1, p2, nrow=1)

##Cool Plot
fancyRpartPlot(fit.rpart$finalModel,
               main = NA,
               sub = "Visualization of Cart1 Trees Influence on pH",
               caption = NA,
               palettes = "PuBuGn",
               type = 2)
fancyRpartPlot(fit.rpart2$finalModel,
               main = NA,
               sub = "Visualization of Cart2 Trees Influence on pH",
               caption = NA,
               palettes = "Purples",
               type = 2)

##ALT PLOT
rpartY <- as.party.rpart(fit.rpart$finalModel)
plot.party(rpartY)
rpartY <- as.party.rpart(fit.rpart2$finalModel)
plot.party(rpartY)
```