---

title: "Group Project"

author: "Group 3 - Yuhui Luo, Adlai Stevenson, Kevin Wong, Kailun Zheng"

date: "4/19/2018"

output: html\_document

---

```{r}

# rm(list = ls())

```

```{r, warning = FALSE, message = FALSE}

# suggested packages

library(MASS)

library(caret)

library(tidyverse)

library(knitr)

library(kableExtra)

library(mlbench)

library(ISLR)

library(ellipse)

library(randomForest)

library(gbm)

library(glmnet)

library(rpart)

library(rpart.plot)

library(klaR)

library(gam)

library(e1071)

# feel free to use additional packages

library(readr)

```

```{r}

credit\_data = read.csv("credit\_data.csv")

credit\_data = credit\_data[-1]

```

```{r}

credit\_data$EDUCATION = ifelse(credit\_data$EDUCATION == 0, 5, credit\_data$EDUCATION)

credit\_data$MARRIAGE = ifelse(credit\_data$MARRIAGE == 0, 3, credit\_data$MARRIAGE)

credit\_data$SEX\_M = ifelse(credit\_data$SEX == 1, 1, 0)

credit\_data$SEX\_F = ifelse(credit\_data$SEX == 2, 1, 0)

credit\_data$EDUCATION = ifelse(credit\_data$EDUCATION == 1, "Graduate",

ifelse(credit\_data$EDUCATION == 2, "College",

ifelse(credit\_data$EDUCATION == 3, "HighSchool",

ifelse(credit\_data$EDUCATION == 4, "Other", "Unknown"))))

credit\_data$MARRIAGE = ifelse(credit\_data$MARRIAGE == 1, "Married",

ifelse(credit\_data$MARRIAGE == 2, "Single", "Other"))

```

```{r}

dmy\_edu = dummyVars(" ~ EDUCATION", data = credit\_data)

new\_edu = predict(dmy\_edu, newdata = credit\_data)

dmy\_mar = dummyVars(" ~ MARRIAGE", data = credit\_data)

new\_mar = predict(dmy\_mar, newdata = credit\_data)

# combine dummied coded variables with the original dataframe

credit\_new = cbind(credit\_data, new\_edu, new\_mar)

# drop original categorical variables and ID

drop\_vars = names(credit\_new) %in% c("ID", "SEX", "EDUCATION", "MARRIAGE")

credit\_new = credit\_new[!drop\_vars]

# head(credit\_new)

```

```{r}

set.seed(3)

cred\_idx = createDataPartition(credit\_new$DEFAULT, p = 0.8, list = FALSE)

cred\_trn = credit\_new[cred\_idx, ]

cred\_tst = credit\_new[-cred\_idx, ]

```

```{r}

accuracy = function(actual, predicted) {

mean(actual == predicted)

}

```