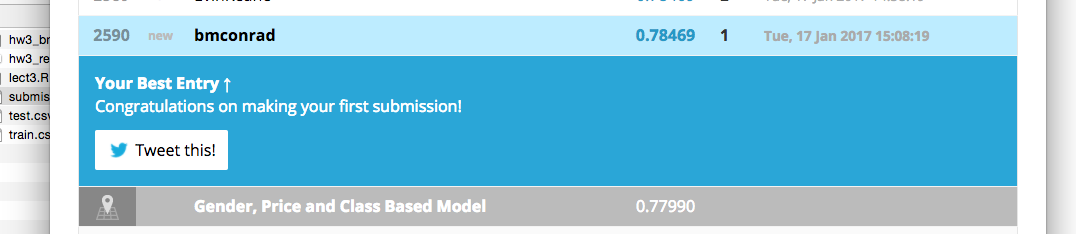
Submission Confirmation:



R Code:

# ---------------------------------------------------------------

# Author: Blake Conrad

# Purpose:

# ~ HW3:

# ~ Register for kaggle titanic competition

# ~ Download the data

# ~ Fit a logistic regression model to the data (glm function)

# ~ Make a submission

# ~ Screenshot the results, and upload with the code.

# ---------------------------------------------------------------

# ---------------------------------------------------------------

# Import the data

# ---------------------------------------------------------------

setwd("/Users/bmc/Desktop/CSCI-49000/week\_2/hw3")

df\_train <- read.csv("train.csv", header=TRUE)

df\_test <- read.csv("test.csv", header=TRUE)

names(df\_train)

names(df\_test)

# ---------------------------------------------------------------

# ---------------------------------------------------------------

# Fit logistic regression model to the data and predict

# ---------------------------------------------------------------

model <- glm(Survived ~ Pclass + Sex\*Age,

data=df\_train,

family="binomial")

summary(model)

df\_test$Survived <- round(predict(model,

newdata = df\_test,

type = "response"))

df\_test$Survived <- round(predict(model, df\_test, type="response"))

df\_test$Survived[is.na(df\_test$Survived)] = 0

write.csv(df\_test[,c("PassengerId","Survived")], "submission.csv", row.names=F)

# ---------------------------------------------------------------

# Put predictions in submission format (PassengerId, Survived)

# ---------------------------------------------------------------

# ---------------------------------------------------------------

Output:

> # ---------------------------------------------------------------

> # Author: Blake Conrad

> # Purpose:

> # ~ HW3:

> # ~ Register for kaggle titanic competition

> # ~ Download the data

> # ~ Fit a logistic regression model to the data (glm function)

> # ~ Make a submission

> # ~ Screenshot the results, and upload with the code.

> # ---------------------------------------------------------------

>

>

> # ---------------------------------------------------------------

> # Import the data

> # ---------------------------------------------------------------

> setwd("/Users/bmc/Desktop/CSCI-49000/week\_2/hw3")

> df\_train <- read.csv("train.csv", header=TRUE)

> df\_test <- read.csv("test.csv", header=TRUE)

> names(df\_train)

[1] "PassengerId" "Survived" "Pclass" "Name" "Sex" "Age" "SibSp"

[8] "Parch" "Ticket" "Fare" "Cabin" "Embarked"

> names(df\_test)

[1] "PassengerId" "Pclass" "Name" "Sex" "Age" "SibSp" "Parch"

[8] "Ticket" "Fare" "Cabin" "Embarked"

> # ---------------------------------------------------------------

>

> # ---------------------------------------------------------------

> # Fit logistic regression model to the data and predict

> # ---------------------------------------------------------------

> model <- glm(Survived ~ Pclass + Sex\*Age,

+ data=df\_train,

+ family="binomial")

> summary(model)

Call:

glm(formula = Survived ~ Pclass + Sex \* Age, family = "binomial",

data = df\_train)

Deviance Residuals:

Min 1Q Median 3Q Max

-2.4792 -0.6683 -0.3740 0.6282 2.5913

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) 4.367487 0.537717 8.122 4.58e-16 \*\*\*

Pclass -1.326218 0.143090 -9.268 < 2e-16 \*\*\*

Sexmale -1.201475 0.435330 -2.760 0.005782 \*\*

Age -0.007696 0.011679 -0.659 0.509906

Sexmale:Age -0.048066 0.014514 -3.312 0.000927 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 964.52 on 713 degrees of freedom

Residual deviance: 635.85 on 709 degrees of freedom

(177 observations deleted due to missingness)

AIC: 645.85

Number of Fisher Scoring iterations: 5

> df\_test$Survived <- round(predict(model,

+ newdata = df\_test,

+ type = "response"))

> df\_test$Survived <- round(predict(model, df\_test, type="response"))

> df\_test$Survived[is.na(df\_test$Survived)] = 0

> write.csv(df\_test[,c("PassengerId","Survived")], "submission.csv", row.names=F)

> # ---------------------------------------------------------------

> # Put predictions in submission format (PassengerId, Survived)

> # ---------------------------------------------------------------

> # ---------------------------------------------------------------