```
library(fastDummies)

ksdf = read.csv("cleaned-ks-data-more.csv", header = TRUE)
data = dummy_cols(ksdf, select_columns = c("category", "launched"))
data = subset(data, select = c(-category, -launched, -X.1, -X, -success_ratio))
train_obs = sample(nrow(data), 30000)
X_train = data[train_obs, ]
X_test = data[-train_obs, ]
Y_train = predict_variable[train_obs]
Y_test = predict_variable[-train_obs]
glm.fits = glm(state01~., data = data, subset=train_obs, family = binomial)
glm.probs = predict(glm.fits,X_test, type='response')
glm.pred = rep(0, nrow(X_test))
glm.pred[glm.probs>0.5] = 1
table(glm.pred, X_test[,'state01'])
```