

Galera,

Ao se cadastrar no Kaggle, display name preencha com FSBDS25_MeuNome.

O script em R abaixo "escora" 0.63636. Minha sugestão é vocês rodarem inicialmente o modelo do jeito que está para conhecer a "sistemática" do Kaggle.

Ao se cadastrarem no Kaggle, incluam no campo "Display name" a tag "FSBDS" antes do nome de vocês, assim ó: FSBDS Diógenes Justo.

Passo 1: Baixem os arquivos train.csv e test.csv para o computador de vocês (se instalação padrão do Windows, na pasta Documentos")

Passo 2: Executem o script abaixo (copiar e colar no R e apertar enter)

Passo 3: Submetam o arquivo predict.csv

A partir daí é trabalhar os dados train e test e também selecionar variáveis diferentes para melhorar o escore de vocês.

```
-----  
-----  
  
# Leitura dos arquivos  
train <- read.csv("train.csv")  
test <- read.csv("test.csv")  
  
# Execução do modelo  
m1<-lm(Survived~Embarked, data=train)  
  
# Previsão dos dados  
p<-predict(m1, newdata=test)  
  
# Classificação de Survived, através do gatilho 0.5  
  
pr <- ifelse(p<0.5, 0, 1)  
  
# Montagem do dataframe para saída de dados  
pr_data <- cbind(test$PassengerId, pr)
```

```
pr_data <- as.data.frame(pr_data)
```

```
# Nomes das colunas
```

```
names(pr_data) <- c("PassengerId", "Survived")
```

```
# Gravação em disco do arquivo a ser submetido no site Kaggle
```

```
write.csv(pr_data, file="predict.csv", row.names = FALSE)
```