

Thank you for accepting the rules.



Titanic: Machine Learning from Disaster

11 months to go

Friday, September 28, 2012

Knowledge • 6,952 teams

Sunday, September 28, 2014

Dashboard

Home

Data

Make a submission

Information

Description

Evaluation

Rules

Prizes

Frequently Asked Questions

Further Reading / Watching

Getting Started With Excel

Getting Started With Python

Getting Started With Rand...

Submission Instructions

Forum

Leaderboard

Visualization

My Team

My Submissions

Forum (189 topics)

Issue when submitting Excel file - column errors?
2 days ago

Basic machine learning course at edX form Caltech
3 days ago

Some help with weka please !!!
4 days ago

Error on loading Excel file
6 days ago

Issue w/ submission for Titanic (newbie)
7 days ago

Struggling with reading csv on Python
12 days ago

Competition Details » Get the Data » Make a submission

Data Files

File Name	Available Formats
train	.csv (59.76 kb)
genderclassmodel	.csv (3.18 kb)
	.py (4.68 kb)
gendermodel	.csv (3.18 kb)
	.py (2.54 kb)
myfirstforest	.csv (3.18 kb)
	.py (3.83 kb)
test	.csv (27.96 kb)

You only need to download one format of each file.

Each has the same contents but use different packaging methods.

VARIABLE DESCRIPTIONS:

```
survival      Survival
               (0 = No; 1 = Yes)

pclass        Passenger Class
               (1 = 1st; 2 = 2nd; 3 = 3rd)

name          Name

sex           Sex

age           Age

sibsp         Number of Siblings/Spouses Aboard

parch         Number of Parents/Children Aboard

ticket        Ticket Number

fare          Passenger Fare

cabin         Cabin

embarked      Port of Embarkation
               (C = Cherbourg; Q = Queenstown; S = Southampton)
```

SPECIAL NOTES:

```
Pclass is a proxy for socio-economic status (SES)
```

1st ~ Upper; 2nd ~ Middle; 3rd ~ Lower

Age is in Years; Fractional if Age less than One (1)

If the Age is Estimated, it is in the form xx.5

With respect to the family relation variables (i.e. sibsp and parch) some relations were ignored. The following are the definitions used for sibsp and parch.

Sibling: Brother, Sister, Stepbrother, or Stepsister of Passenger Aboard Titanic

Spouse: Husband or Wife of Passenger Aboard Titanic (Mistresses and Fiances Ignored)

Parent: Mother or Father of Passenger Aboard Titanic

Child: Son, Daughter, Stepson, or Stepdaughter of Passenger Aboard Titanic

Other family relatives excluded from this study include cousins, nephews/nieces, aunts/uncles, and in-laws. Some children travelled only with a nanny, therefore parch=0 for them. As well, some travelled with very close friends or neighbors in a village, however, the definitions do not support such relations.