

Making competitions classifiers

Competition organization

Judges

Give training set

Judges training dataset

Competitors know the labels

Withhold test set

Judges test dataset

Competitors do not know the labels

Submissions management

Submission logs

Score calculation

Verification

Data analysis

Summaries

Distribution charts

Mosaic plots

etc.

Preliminary data transformations

Categorical variables handling

Feature extraction

Dimension reduction

etc.

Study

Optional

Come up with relevant data transformations

Come up with different feature engineering procedures

Classifier training

Do multiple times

Split the ***judges training dataset***

Training dataset

Test dataset

Train classifiers

Different data transformations

Different classifier algorithms

Classifier measurements

Accuracy, Precision, Recall

Confusion matrix

ROC

etc.

Pick the best classifier and transformation

Make the competition classifier

1. Transform the ***judges training dataset*** according to findings in the study

2. Train a classifier with the method found to be best

3. Transform ***judges test dataset***

4. Run the classifier over the ***transformed judges test dataset***

5. Submit the found labels

Submission preparation