

Train/Test Card ++

When the training set does not match the test set, several solutions exist.

- It is possible to use the information common to the training and test sets to create two data sets with the same variables. This allows you to train the algorithm on the training set and apply it to the problem at hand.
- If you have the intrinsic belief that the data does not behave the same between the two datasets, you will need to collect some additional variables or collect a whole new dataset.
- There are statistical solutions to fit the former data to the latter, if the difference is small.