Cross-Validation Approaches for Replicability in Psychology

Introduction

"The ability to replicate a scientific discovery or finding is one of the features that distinguishes science from non-science and pseudo-science "

Crisis:

A systematic study designed to assess the reproducibility of psychological science (Open Science Collaboration, 2015) found that the mean effect size of the replicated studies was half of that of the originally conducted studies, and, even more strikingly, only 36% of replication studies had statistically significant results.

strategies

- (i) improvement in research methodologies;
- (ii) promotion of replication attempts.

Simulating replicability via cross-validation techniques

Advantages:

- Cross- validation entails a set of techniques that partition the dataset and repeatedly generate models and test their future predictive power (Browne, 2000).

- Cross-validation has the computational advantage that it avoids fitting a model too closely to the peculiarities of a data set (overfitting).

Holdout Cross-Validation

The portion of data used for the training dataset is randomly selected, and the remaining part of the data, generally a fraction of 1/3 of the data, is assigned to the testing dataset.

For instance, if collected data consists of 1,000 observations, a training set would typically consist of 667 observations, while 333 observations would be kept aside as a testing dataset for verifying the generated model.

k-fold Cross-Validation

This procedure is repeated k-times, each time selecting a different fold for testing purposes and the other folds (k-1) as training dataset.

- This method avoids the randomness emanating from estimates produced by splitting the data only once.
- The disadvantage is that the training algorithm has to be rerun from scratch k times

Leave-One-Subject-Out Cross-Validation

Leave-One-Trial-Out Cross-Validation

Prediction Error

RMSE, R-square

It's better to combine them with model fitting index

Software

Scikit-learn (Pedregosa et al., 2011), Pylearn2 (Goodfellow et al., 2013), PyMVPA (Hanke et al., 2009), Statistics and Machine learning toolbox (Matlab; MathWorks, Natick, MA), e1071 (Meyer et al., 2017), caret (Kuhn, 2008), and Microsoft Azure (Microsoft Corporation)

- Return

Software

Scikit-learn (Pedregosa et al., 2011), Pylearn2 (Goodfellow et al., 2013), PyMVPA (Hanke et al., 2009), Statistics and Machine learning toolbox (Matlab; MathWorks, Natick, MA), e1071 (Meyer et al., 2017), caret (Kuhn, 2008), and Microsoft Azure (Microsoft Corporation)