

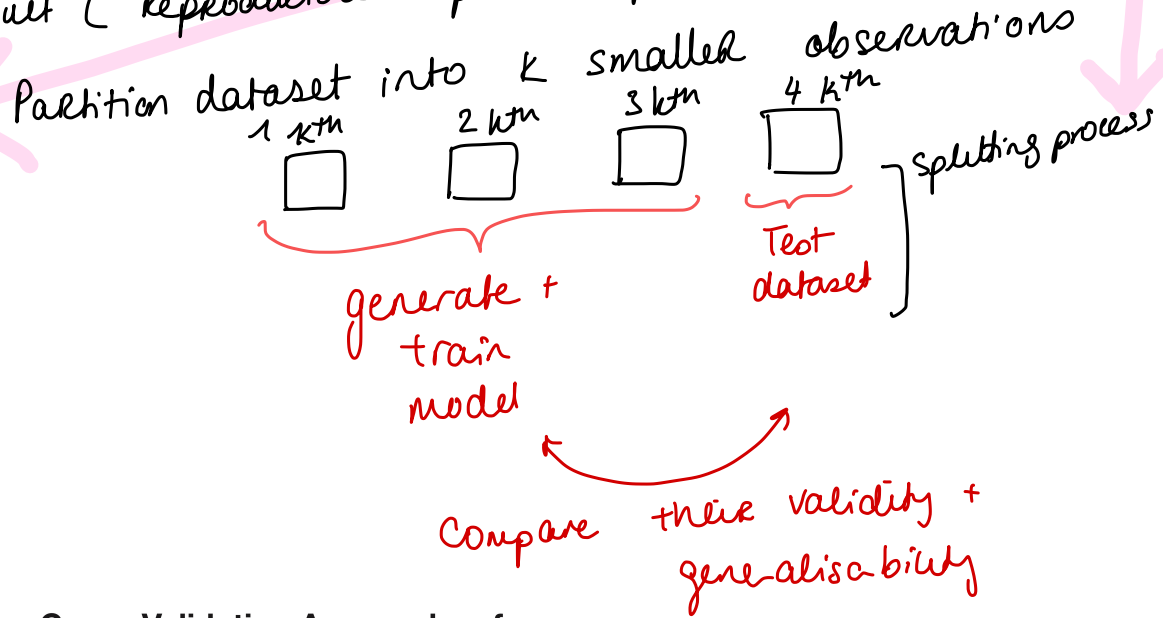
Statistics > Applied Statistics > Stat Analysis > Conclusion > Take sample > Manipulate

Cross Validation

the idea

We need to convince ourselves that 1st conclusion \neq merely isolated event. That our cross validation produces the same result (reproducible upon replication)

the method



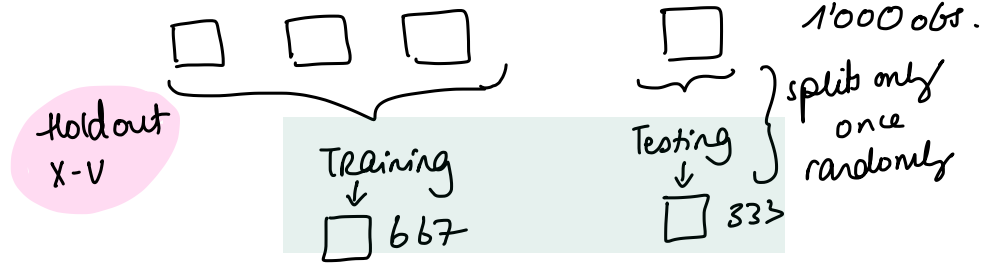
ref :
Koul, 2018

Cross-Validation Approaches for Replicability in Psychology

Atesh Koul¹, Cristina Becchio^{1,2} and Andrea Cavallo^{1,2*}

¹ C'IMON, Cognition, Motion and Neuroscience Unit, Fondazione Istituto Italiano di Tecnologia, Genova, Italy; ² Department of Psychology, University of Torino, Torino, Italy

creds: andreytyeo@gmail.com



X-V = cross-validation

- + ↓ computation load
- ↑ variance

+ advantage

- disadvantage

K-fold X-V



four equal folds of observation



1 fold = Testing

3 folds (k-1 folds) = Training-
accuracy = A_1

Repeated k times generate k accuracy: $\{A_1, A_2, A_3, A_4\}$
+ avoids randomness from merely having 1 fold

- + lower variance
- training data has to be run from scratch k times which means it takes k times as much computation to make an evaluation.

leave 1 subject out X-V
decide ? = k folds

↓
randomly split into



until all subjects have been part of both testing dataset. (like a cross-over clinical trial)
used to predict clinical status of new patient

X-V = cross-validation

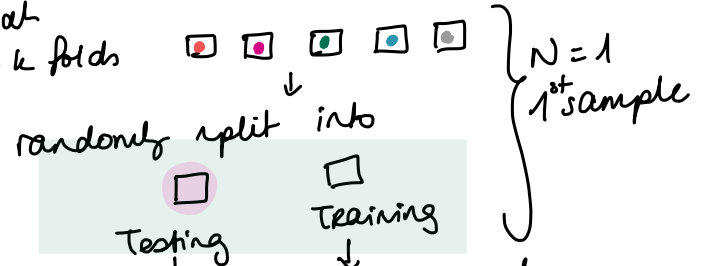
+ advantage

- disadvantage

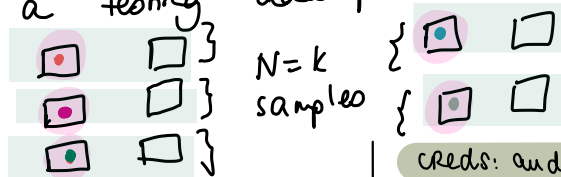
leave one trial out

k = N trials

decide ? = k folds



until all samples have been used as a testing data point



creds: andeeytyeo@gmail.com