Applicability of Cross-validation Error in Neural Networks

compiled by D.Gueorguiev, 7/28/24

Cross-validation is a technique to measure the predictive performance of a model. This document discusses different methods of cross-validation and their applicability to neural networks.

Holdout Sample

Data is split into two groups – training and test data. The training set is used to train the learner. The test set is used to estimate the error rate of the trained model. This method has two basic drawbacks – in a sparse/small data set, one may not be able to set aside a reasonable portion of the data for testing. Since it is a single repetition of the train-and-test experiment, the error estimate is not stable. If the data sample has a non-representative (“bad”) split, the estimate will not be reliable.

Three-Way Split

The available data is partitioned into three sets: training, validation, and test set. The prediction model is trained on the training set and is evaluated on the validation set. For example, in case of a neural network, the training set is used to find the optimal weights with the back-prop rule. The validation set may be used to find the optimum number of hidden layers or to determine a stopping rule for the back-prop algorithm. Training and validation may be iterated until a ‘best’ model is selected. The final model is assessed using the test set.

A typical split is 50% for training data, and 25% each for validation and test set.

With three-way split, the model selection, and the true error rate computation can be carried out simultaneously. The error rate estimate of the final model on validation data will be biased (smaller than the error rate) since the validation set is used to select the final model. Hence a third independent part of the data, the test data, is required.

After assessing the final model on the test set, the model must not be fine-tuned any further.

Note: data insufficiency often does not allow three-way split. The limitation of the holdout or the three-way split can be overcome with a family of resampling methods at the expense of higher computational cost.

Cross-Validation

References

[Applied Data Mining and Statistical Learning, Lecture 2, STAT 897D, Penn State](https://online.stat.psu.edu/stat857/node/161/)

[Cross-validatory Choice and Assessment of Statistical Prediction, M. Stone, University College London, 1974](https://github.com/dimitarpg13/probabilistic_machine_learning/blob/main/applied_statistics/articles/Cross-validatory_Choice_and_Assessment_of_Statistical_Predictions_Stone1974.pdf)

[How to use k-fold cross validation error in neural network, StackOverflow, last modified 5/2017](https://stackoverflow.com/questions/25889637/how-to-use-k-fold-cross-validation-in-a-neural-network)