2025/06/27 03:22

A validation dataset is a dataset of examples used to tune the hyperparameters of a classifier. It is sometimes also called the development set or the "dev set". In artificial neural networks, a hyperparameter is, for example, the number of hidden units. It, as well as the testing set, should follow the same probability distribution as the training dataset.

In order to avoid overfitting, when any classification parameter needs to be adjusted, it is necessary to have a validation dataset in addition to the training and test datasets. For example, if the most suitable classifier for the problem is sought, the training dataset is used to train the candidate algorithms, the validation dataset is used to compare their performances and decide which one to take and, finally, the test dataset is used to obtain[citation needed] the performance characteristics such as accuracy, sensitivity, specificity, F-measure, and so on. The validation dataset functions as a hybrid: it is training data used by testing, but neither as part of the low-level training nor as part of the final testing[citation needed].

The basic process of using a validation dataset for model selection (as part of training dataset, validation dataset, and test dataset) is:

Since our goal is to find the network having the best performance on new data, the simplest approach to the comparison of different networks is to evaluate the error function using data which is independent of that used for training. Various networks are trained by minimization of an appropriate error function defined with respect to a training data set. The performance of the networks is then compared by evaluating the error function using an independent validation set, and the network having the smallest error with respect to the validation set is selected. This approach is called the hold out method. Since this procedure can itself lead to some overfitting to the validation set, the performance of the selected network should be confirmed by measuring its performance on a third independent set of data called a test set.

An application of this process is in early stopping, where the candidate models are successive iterations of the same network, and training stops when the error on the validation set grows, choosing the previous model (the one with minimum error).

From: https://neurosurgerywiki.com/wiki/ - **Neurosurgery Wiki**

Permanent link: https://neurosurgerywiki.com/wiki/doku.php?id=validation_dataset



Last update: 2024/06/07 02:59