

Train/test split involves splitting the dataset into training and testing sets respectively, which are mutually exclusive.

After which, you train with the training set and test with the testing set.

This will provide a more accurate evaluation on out-of-sample accuracy because the testing dataset is not part of the dataset that has been used to train the data. It is more realistic for real-world problems. This means that we know the outcome of each data point in the dataset, making it great to test with.

Since this data has not been used to train the model,

the model has no knowledge of the outcome of these data points.

So, in essence, it's truly out-of-sample testing.

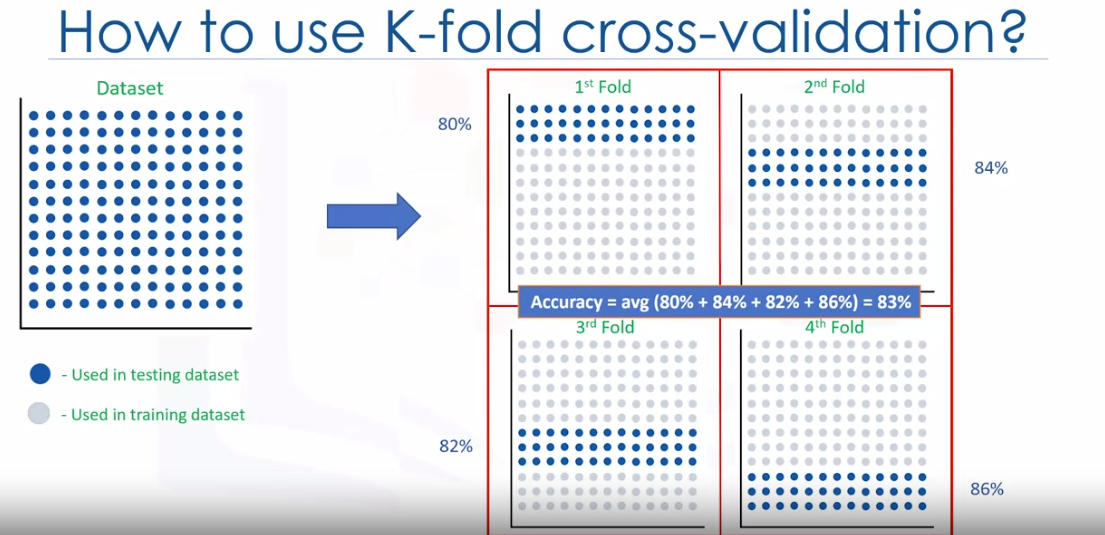
However, please ensure that you train your model with the testing set afterwards,

as you don't want to lose potentially valuable data.

But it still has some problems due to this dependency. Another evaluation model, called K-fold cross-validation, resolves most of these issues.

How do you fix a high variation that results from a dependency?

Well, you average it.



The entire dataset is represented by the points in the image at the top left. If we have K equals 4 folds,

then we split up this dataset as shown here.

In the first fold for example, we use the first 25 percent of the dataset for testing and the rest for training.

The model is built using the training set and is evaluated using the test set. Then, in the next round or in the second fold, the second 25 percent of the dataset is used for testing and the rest for training the model.

Again, the accuracy of the model is calculated.

We continue for all folds. Finally, the result of all four evaluations are averaged.

That is, the accuracy of each fold is then averaged,

keeping in mind that each fold is distinct,

where no training data in one fold is used in another. K-fold cross-validation in its simplest form performs multiple train/test splits, using the same dataset where each split is different.

Then, the result is average to produce a more consistent out-of-sample accuracy.