



This PDF is generated from authoritative online content, and is provided for convenience only. This PDF cannot be used for legal purposes. For authoritative understanding of what is and is not supported, always use the online content. To copy code samples, always use the online content.

Genesys Knowledge Management User Guide

Cross-Validation

4/25/2025

Cross-Validation

This topic describes part of the functionality of **Genesys Content Analyzer**.

In cross-validation, Training Server follows these steps:

1. It builds one model using all of the data.
2. It divides the data into x partitions, where $x = 3, 5$, or 10 .
3. It builds a number of partial models: as many as there are partitions, each one using a different combination of $x - 1$ partitions.

For example, if the data is divided into the three partitions A, B, and C, Training Server builds model X using partitions A and B, model Y using partitions A and C, and model Z using partitions B and C.

4. It tests each of these partial models against the partition that it omitted when it was built.

In the example, it tests model X against partition C, model Y against partition B, and model Z against partition A.

5. It aggregates the results of all these tests and presents them as the rating of the entire model.

These ideas underlie the concept of cross-validation:

- The best way to test a model is to apply it to data that was not used in building the model.
- A model built using most of the data is usefully similar to the model built using all of the data, so the results of testing (for example) all possible 90-percent models are a good indication of the quality of the 100-percent model.

Because cross-validation adds to the time required to build a model, you may not want to select cross-validation for very large training objects or for objects for which you selected training quality level 6.