



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.

Content Analyzer Plug-in for GAX

[Schedule Training](#)

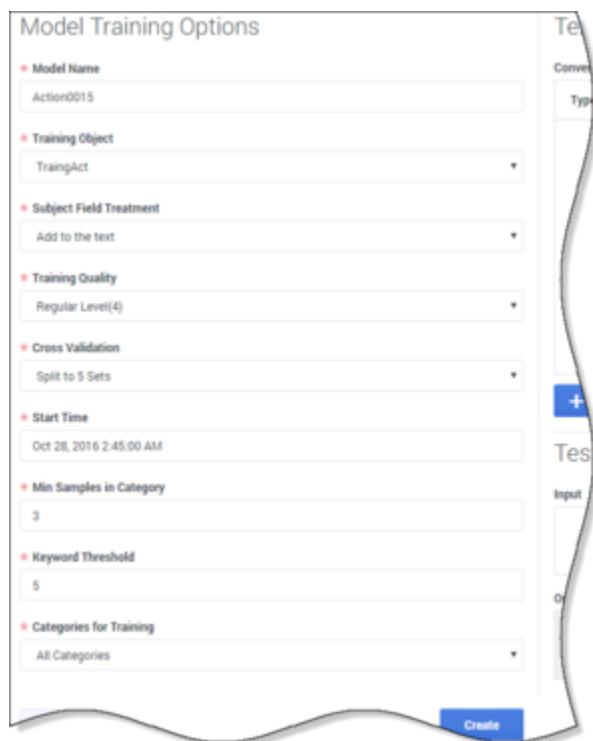
12/17/2025

Schedule Training

On the **Training Schedule** tab, set up a training session.

- To populate fields with the values of an existing training session, click the clone  icon. This is a handy way to change the time that an existing session is scheduled to run.
- To populate fields with default values, click the plus-sign icon.

In either case, the **Model Training Options** dialog appears.



Model Training Options

1. **Model names** must be no more than 21 characters long, and must use only the **allowed characters**.
2. **Subject Field Treatment**
 - **Ignore**—Training does not consider the content of the **Subject** field.
 - **Add to the text**—Training considers the content of the **Subject** field.
 - **Add with double weight**—Training gives the content of the **Subject** field twice as much importance as the content of the e-mail body.
3. **Training Quality**—If you know that the Training Data Object contains many wrongly categorized text objects, use Unreliable Levels 10-12. Otherwise, use Draft, or Regular Levels 1-6. Note the following:

- The Regular Levels and the Unreliable Levels form two independent scales that are not easily comparable. Within each, a higher number means better quality. The only way to know for sure whether, for example, Unreliable Level (11) will produce better or worse results than Regular Level (4), is to create one model with each setting and test them.
 - These levels actually determine the number of words that the system considers and the number of iterations the training process runs. Increasing both of those should increase the quality of the resulting model, but at higher levels it may not. Again, the only way to know is to test the resulting models, preferably with **cross-validation**.
 - Training time increases as you move from Draft quality to Regular Level 3 quality. But once the quality goes above 3, there is not much difference in training time. Genesys recommends that you use the lowest quality only when you want to obtain a preliminary reading of the model's quality estimation. For production, use quality 2-6.
4. **Cross-Validation** is explained on a separate page. Select either no cross-validation, or cross-validation that splits the data into 3, 6, or 10 sets. If you select cross-validation, training produces an accuracy rating for the model along with the model itself. This has the advantage of not requiring an extra testing step, but it increases the training time.
 5. For **Start Time**, because training can use a large proportion of system resources, you will probably want to schedule it for nonpeak hours. Be sure to set a time later than the present moment.
 6. **Min Samples in Category** is the minimum number of text objects that a category must have in order to be included in training. Categories with no or few text objects make poor subjects for training.
 7. **Keyword Threshold** is the minimum number of text objects that a keyword must occur in for that keyword to be considered in training. A relatively high value for this setting can reduce training time, but it can also reduce quality. What counts as a high or low value for this setting depends on the total size of the Training Data Object. For example, if a Training Data Object has 5 to 10 text objects per category, a high keyword threshold might be **2** or **3**. If a Training Data Object has 30 to 50 text objects per category, a high keyword threshold might be **20**.
 8. **Categories for Training** is **All Categories** or **Terminal Categories Only**. A "terminal category" is one that contains no subcategories. It may be that a category tree uses nonterminal categories mostly for organizing the terminal categories. If so, few or no text objects are associated with the nonterminal categories, and there is little to be gained by including the nonterminal categories in training.
 9. You can clean up your Training Data Object by using the **Text Preprocessing pane** (to the right of **Model Training Options**) to remove extraneous text.