



String Evaluators

A string evaluator is a component within LangChain designed to assess the performance of a language model by comparing its generated outputs (predictions) to a reference string or an input. This comparison is a crucial step in the evaluation of language models, providing a measure of the accuracy or quality of the generated text.

In practice, string evaluators are typically used to evaluate a predicted string against a given input, such as a question or a prompt. Often, a reference label or context string is provided to define what a correct or ideal response would look like. These evaluators can be customized to tailor the evaluation process to fit your application's specific requirements.

To create a custom string evaluator, inherit from the `StringEvaluator` class and implement the `_evaluate_strings` method. If you require asynchronous support, also implement the `_aevaluate_strings` method.

Here's a summary of the key attributes and methods associated with a string evaluator:

- `evaluation_name`: Specifies the name of the evaluation.
- `requires_input`: Boolean attribute that indicates whether the evaluator requires an input string. If True, the evaluator will raise an error when the input isn't provided. If False, a warning will be logged if an input *is* provided, indicating that it will not be considered in the evaluation.
- `requires_reference`: Boolean attribute specifying whether the evaluator requires a reference label. If True, the evaluator will raise an error when the reference isn't provided. If False, a warning will be logged if a reference *is* provided, indicating that it will not be considered in the evaluation.

String evaluators also implement the following methods:

- `aevaluate_strings`: Asynchronously evaluates the output of the Chain or Language Model, with support for optional input and label.
- `evaluate_strings`: Synchronously evaluates the output of the Chain or Language Model, with support for optional input and label.

The following sections provide detailed information on available string evaluator implementations as well as how to create a custom string evaluator.



Criteria Evaluation

In scenarios where you wish to assess a model's output using a specific rubric or criteria set, the criter...



Custom String Evaluator

You can make your own custom string evaluators by inheriting from the StringEvaluator class and impl...



Embedding Distance

To measure semantic similarity (or dissimilarity) between a prediction and a reference label string, you ...



String Distance

One of the simplest ways to compare an LLM or chain's string output against a reference label is by us...