



Comparison Evaluators

Comparison evaluators in LangChain help measure two different chain or LLM outputs. These evaluators are helpful for comparative analyses, such as A/B testing between two language models, or comparing different versions of the same model. They can also be useful for things like generating preference scores for ai-assisted reinforcement learning.

These evaluators inherit from the `PairwiseStringEvaluator` class, providing a comparison interface for two strings - typically, the outputs from two different prompts or models, or two versions of the same model. In essence, a comparison evaluator performs an evaluation on a pair of strings and returns a dictionary containing the evaluation score and other relevant details.

To create a custom comparison evaluator, inherit from the `PairwiseStringEvaluator` class and overwrite the `_evaluate_string_pairs` method. If you require asynchronous evaluation, also overwrite the `_aevaluate_string_pairs` method.

Here's a summary of the key methods and properties of a comparison evaluator:

- `evaluate_string_pairs`: Evaluate the output string pairs. This function should be overwritten when creating custom evaluators.
- `aevaluate_string_pairs`: Asynchronously evaluate the output string pairs. This function should be overwritten for asynchronous evaluation.
- `requires_input`: This property indicates whether this evaluator requires an input string.
- `requires_reference`: This property specifies whether this evaluator requires a reference label.

Detailed information about creating custom evaluators and the available built-in comparison evaluators are provided in the following sections.



Custom Pairwise Evaluator

You can make your own pairwise string evaluators by inheriting from `PairwiseStringEvaluator` class and...



Pairwise Embedding Distance

One way to measure the similarity (or dissimilarity) between two predictions on a shared or similar inp...



Pairwise String Comparison

Often you will want to compare predictions of an LLM, Chain, or Agent for a given input. The StringCo...