

Knowledge Graph Dataset

In this section, we filter the main source of data creation.

> (ie: Wikipedia for FactBench Dataset)

We use a strategy to check for similarity, and if the similarity percentage falls below a certain threshold, we remove that data.

> However, in some cases, removing the data leaves us with no input data.

When this happens, we stop using the strategy altogether.

When two of the language models believe the answer is correct, and two believe it is wrong, we switch to a more advanced model, such as a commercial one or a model with more parameters.

This model is then used to make the final decision.

