

Popular RAG Evaluation Frameworks

Instructor

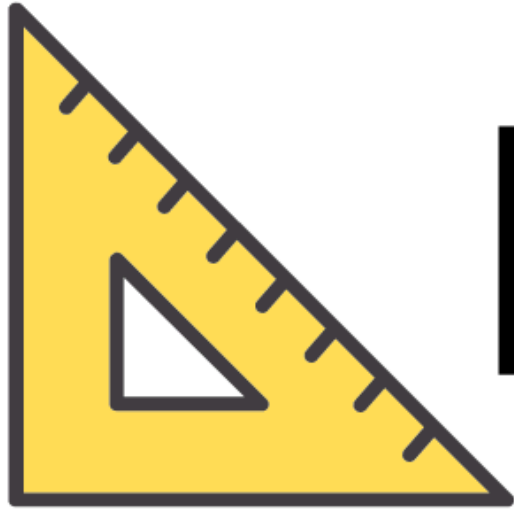
Dipanjan Sarkar

Head of Community & Principal AI Scientist at Analytics Vidhya

Google Developer Expert - ML & Cloud Champion Innovator

Published Author



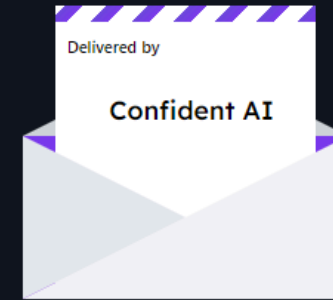


ragas

DeepEval.

\$ the open-source LLM evaluation framework

Get Started





Why DeepEval?

- **Comprehensive LLM Evaluation Metrics:**
 - Offers a variety of evaluation metrics such as G-Eval, RAGAS, faithfulness, contextual recall, and hallucination metrics.
 - Enables robust and tailored evaluation of LLM outputs for specific use cases.
- **Flexible Deployment Options:**
 - Supports both local machine evaluations and cloud-based evaluations on Confident AI's infrastructure.
 - Provides flexibility for different environments and resource constraints.
- **Efficient Bulk Evaluation:**
 - Allows evaluation of entire datasets in parallel with minimal code (less than 20 lines).
 - Includes CLI or the `evaluate()` function to streamline large-scale LLM evaluations.

Why DeepEval?

- **Seamless CI/CD Integration:**
 - Compatible with any CI/CD pipeline.
 - Enables automated, continuous LLM evaluation across development cycles with real-time production-ready evaluations.
- **Customizable and Extensible:**
 - Provides a base metric class for creating custom metrics.
 - Tailors evaluations to unique RAG (Retrieval-Augmented Generation) or fine-tuning requirements.

Thank You
