ML Flow Evaluation Metrics:

semantic_similarity_metric

- Measures meaning similarity using TF-IDF.
- Target: Values close to 10/100% (9/90%+). Indicates perfect semantic alignment.

readability_improvement_metric

- Measures improvement in sentence simplicity and structure.
- Target: positive values/percentages – model makes text more readable.

grammar_quality_score/llm_judge_metric

- LLM assigns a score from 0-10 based on corrected grammar quality.
- Target: 10/100% – perfect grammar with no issues.