RAG Evaluation Report

Generated: 9/25/2025, 1:04:37 PM
Course ID: 36
Retrieval @k: 5

Basic Performance

Total Questions: 40
Successful Responses: 40
Response Rate: 100.00%
Average Answer Length: 200 chars

Core RAG Metrics

Context Precision: 85.5%
Context Recall: 95.4%
Answer Correctness: 30.9%
Faithfulness: 67.9%
Answer Relevance: 23.7%
Semantic Similarity: 36.3%

Quality Metrics

Factual Accuracy: 21.7%
Completeness: 26.3%

NLP Evaluation Metrics

BLEU-1: 44.4%
BLEU-4: 24.2%
ROUGE-1: 96.4%
ROUGE-2: 67.5%
ROUGE-L: 43.6%
METEOR: 27.6%

BERT Precision: 44.4%
BERT Recall: 96.4%
BERT F1: 55.2%

Overall Assessment

Overall RAG Quality Score: 56.6%

Performance Grade: Good