

Multi-Turn RAG Benchmark Report

Generated: 2025-06-27 19:03

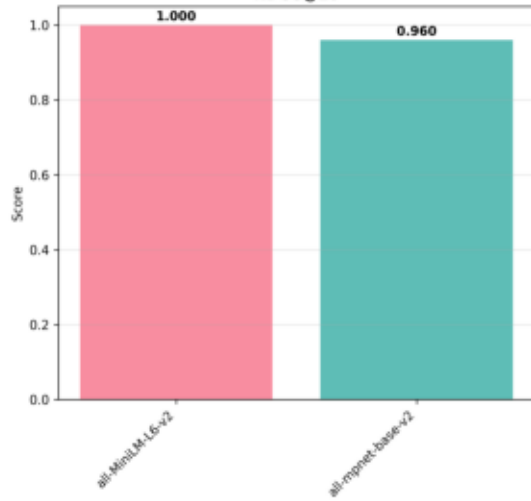
Models Tested: 2

Datasets: 1

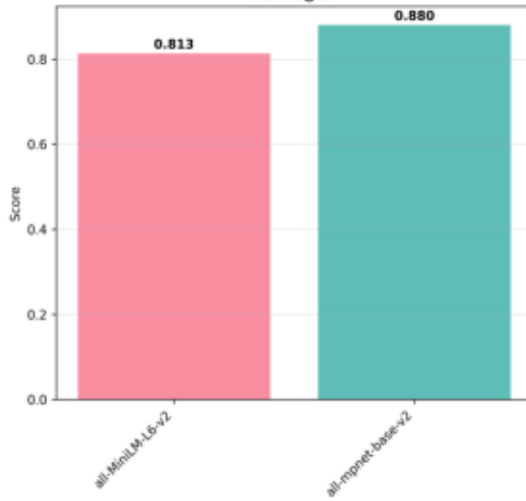
Total Evaluations: 2

Performance Overview

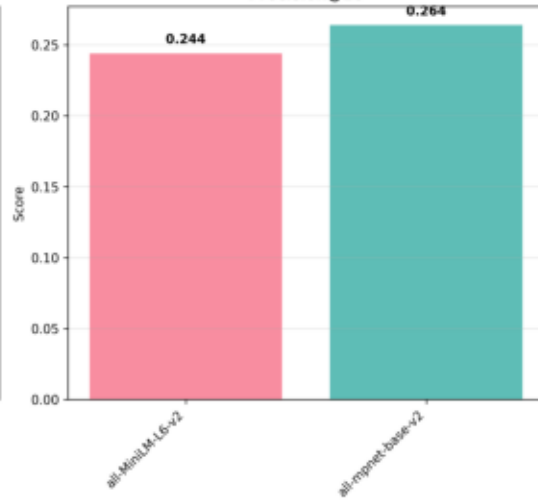
NDCG@10



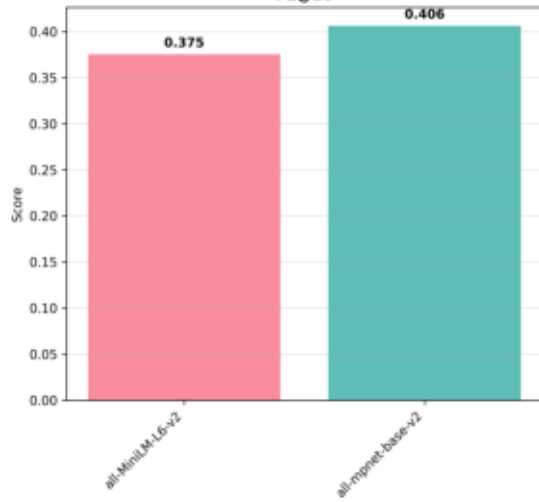
Recall@10



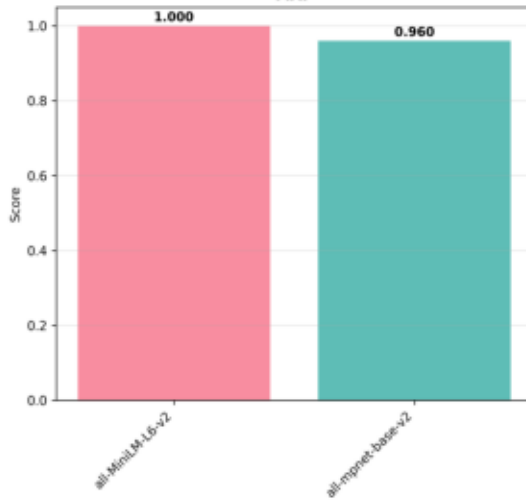
Precision@10



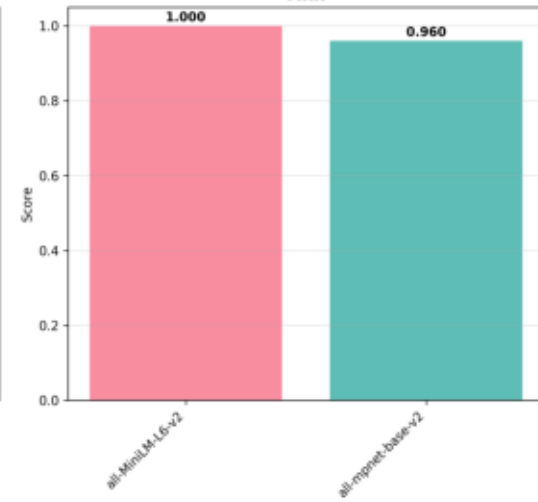
F1@10



MAP

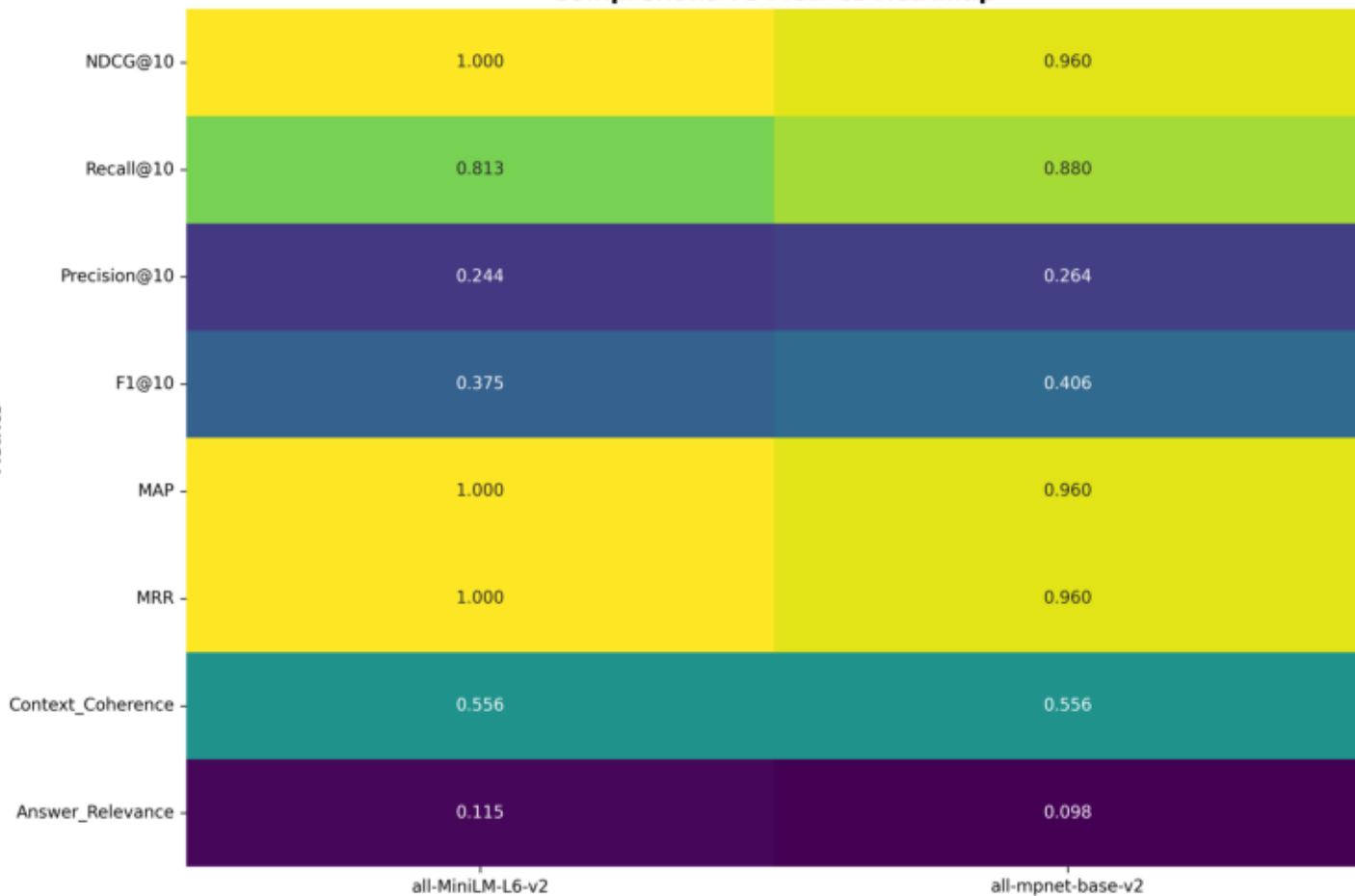


MRR



Comprehensive Metrics Heatmap

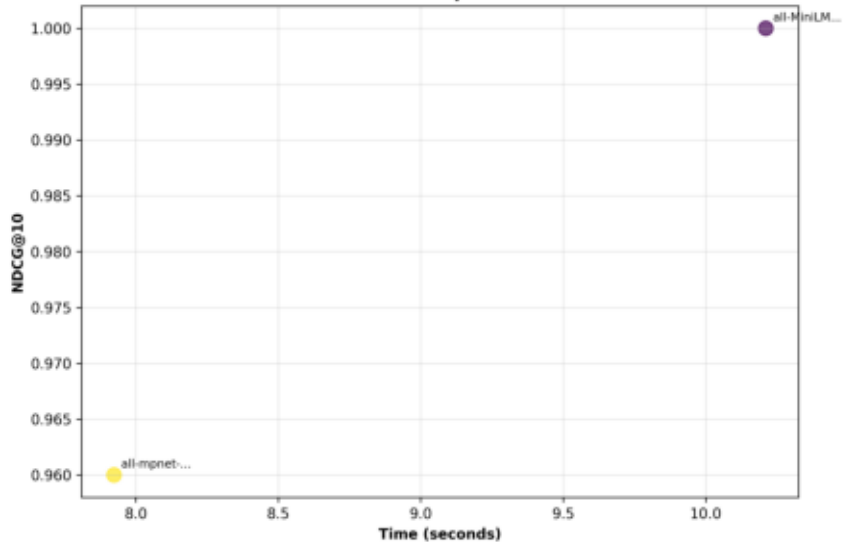
Metrics



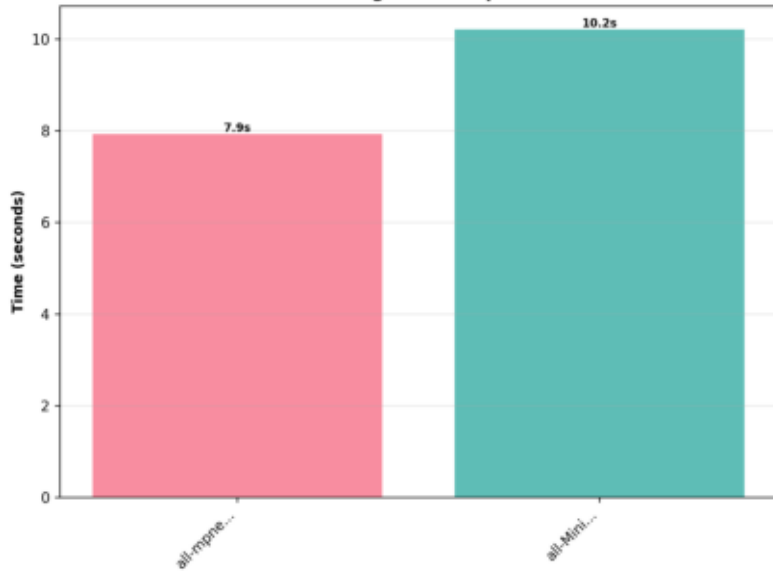
Models

Score

Performance vs Speed Trade-off

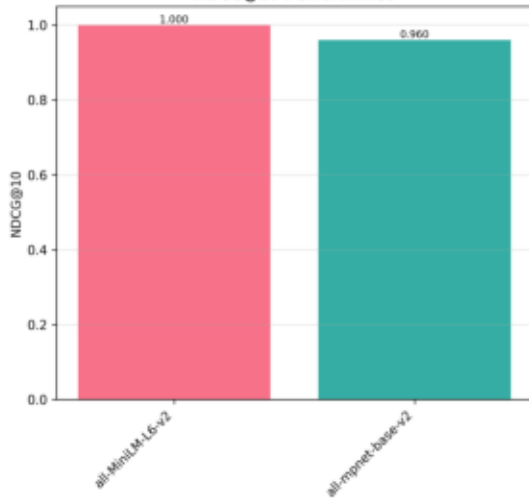


Processing Time Comparison

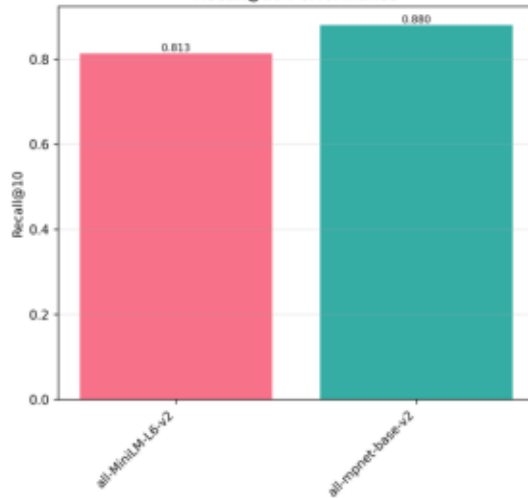


Individual Metric Performance Comparison

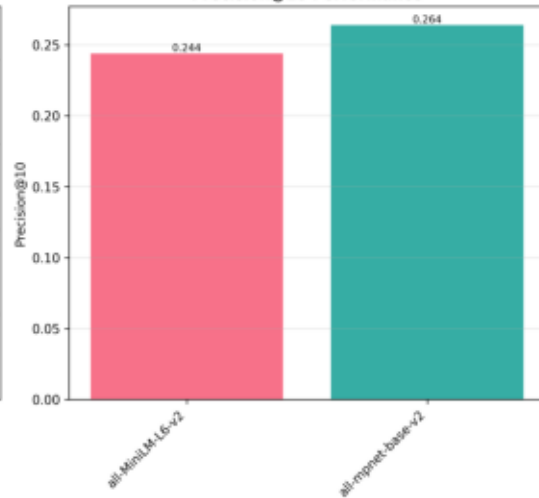
NDCG@10 Performance



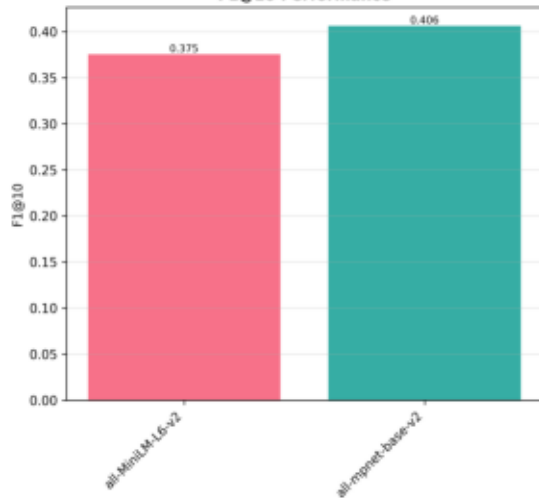
Recall@10 Performance



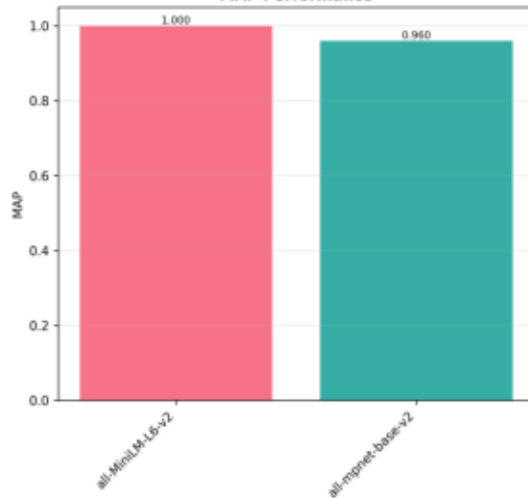
Precision@10 Performance



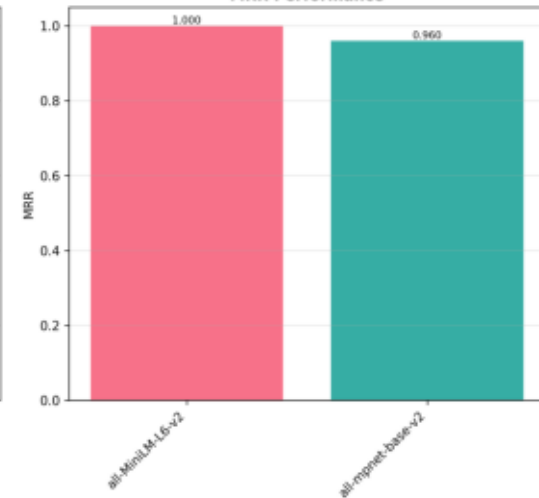
F1@10 Performance



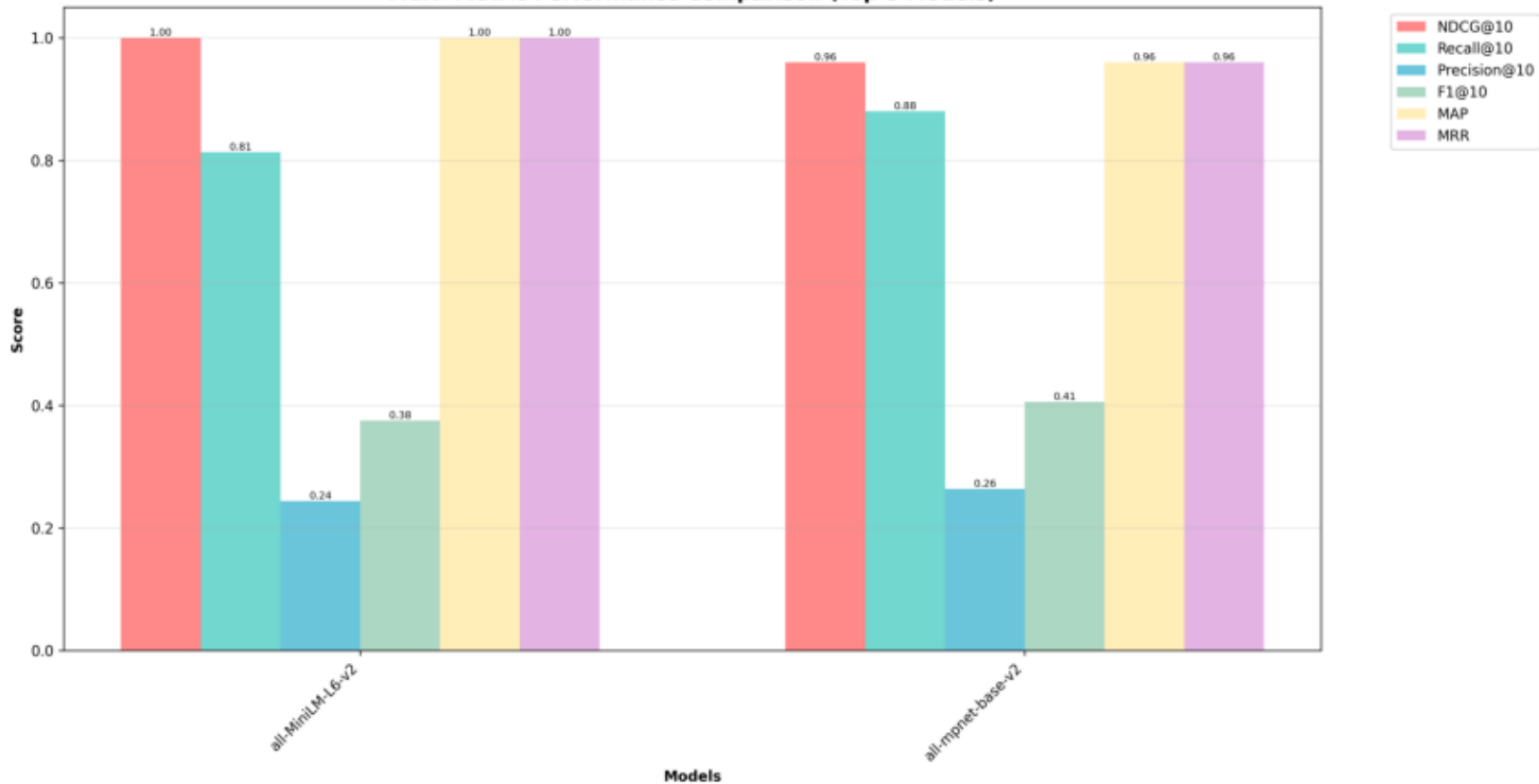
MAP Performance



MRR Performance



Multi-Metric Performance Comparison (Top 8 Models)



Top Performers by Metric

