

Benchmark: MMLU Abstract Algebra
Performance Metric: Accuracy
Correlation Metric: Spearman

