

Benchmark: MMLU Conceptual Physics  
Performance Metric: Accuracy  
Correlation Metric: Kendall

