

Benchmark: MMLUHS US History  
Performance Metric: Accuracy  
Correlation Metric: Kendall

