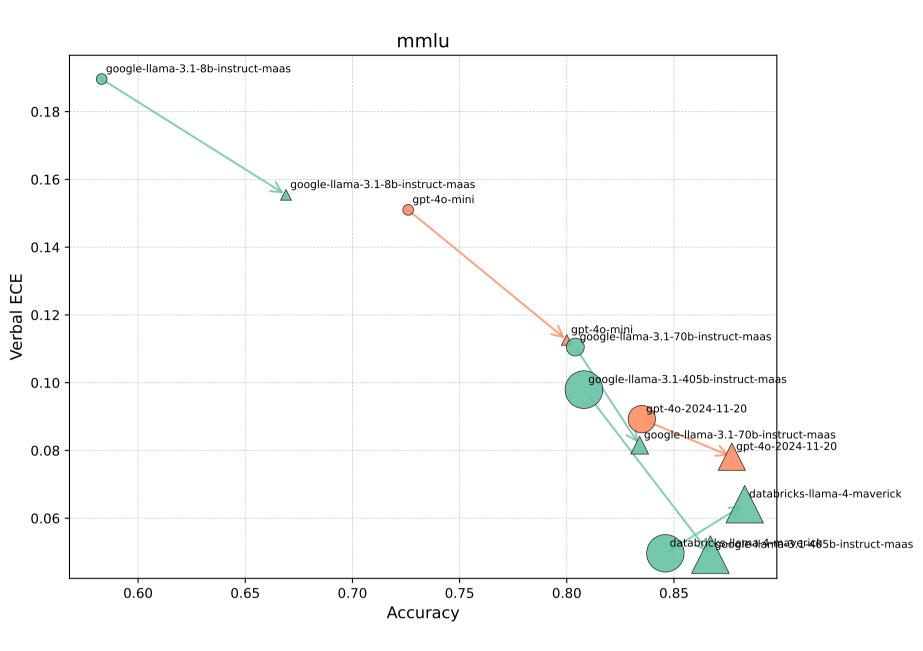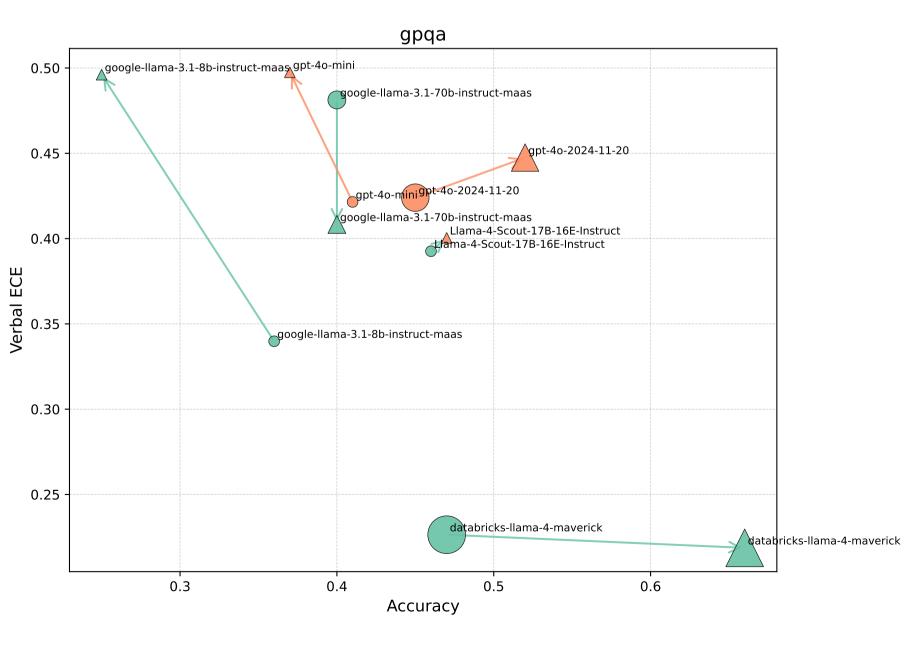Model Performance by Benchmark