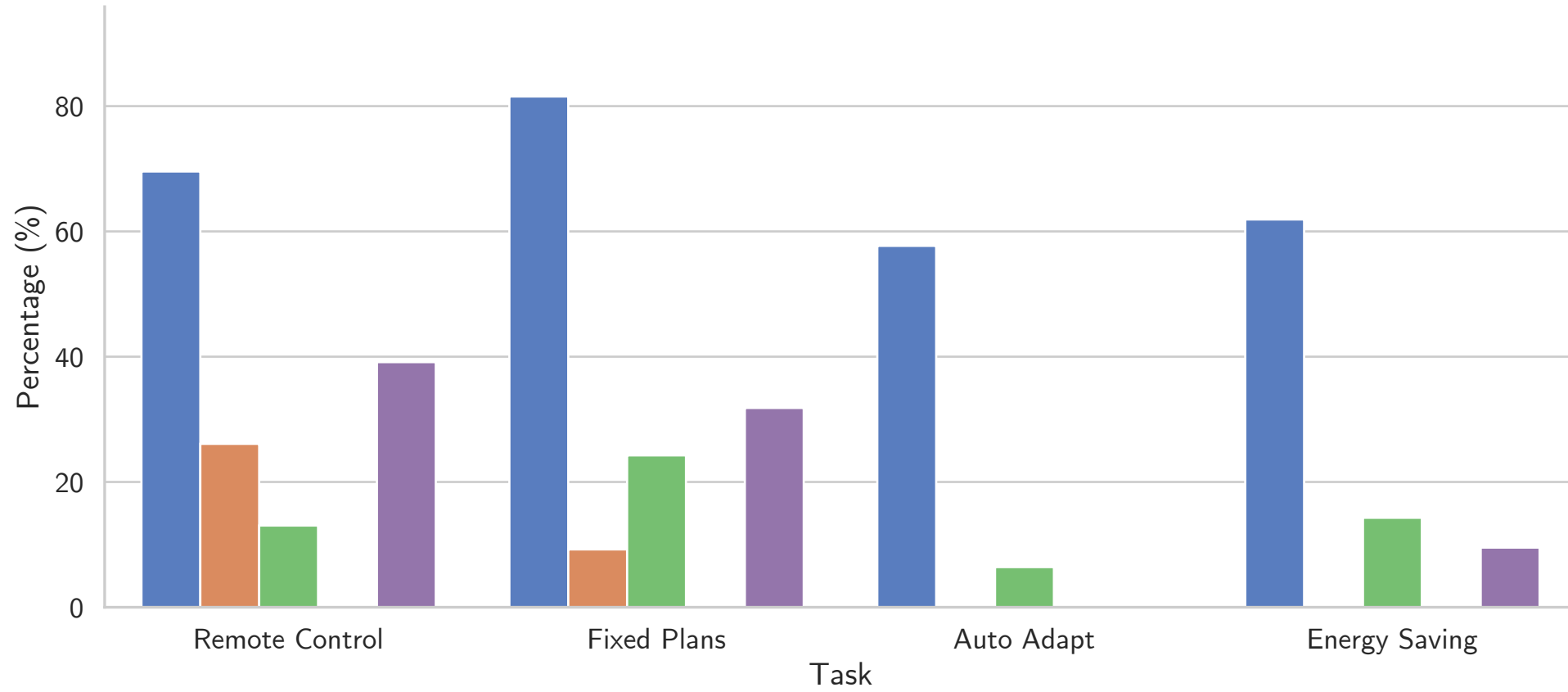
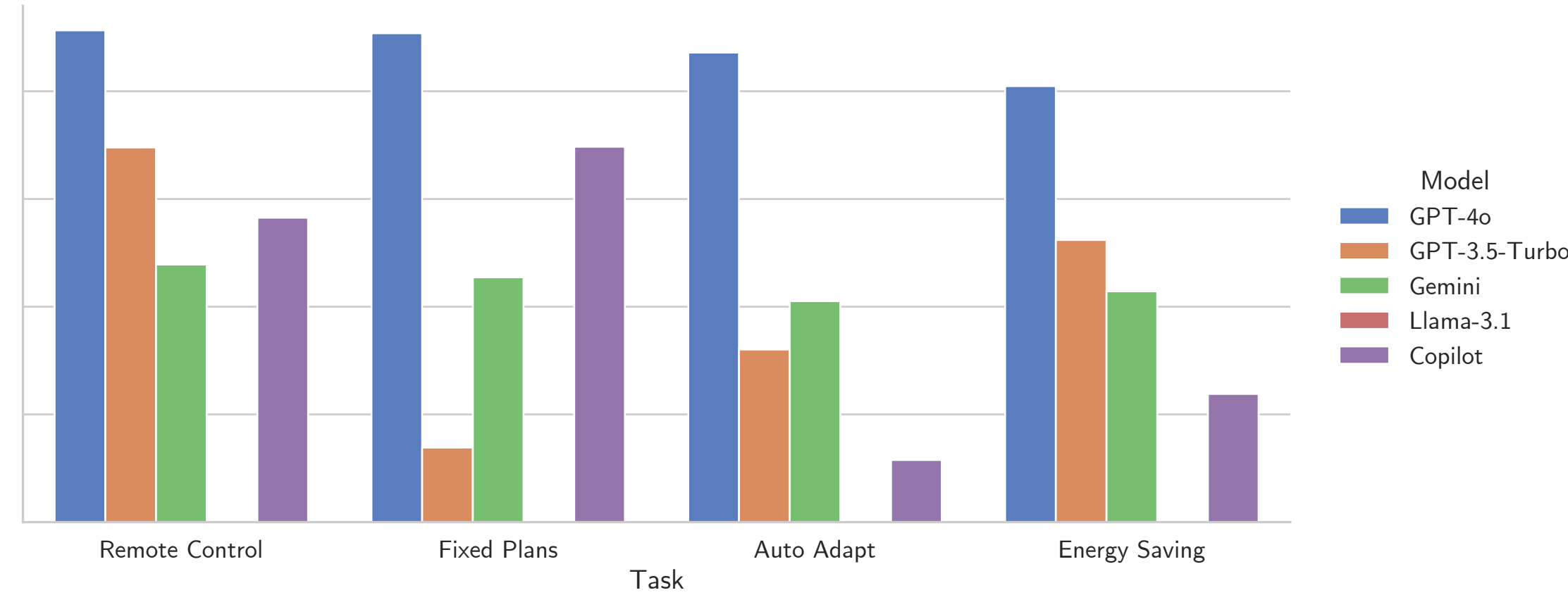


PASS RATE by Task



COMPILABLE RATE by Task



- Model
- GPT-4o
 - GPT-3.5-Turbo
 - Gemini
 - Llama-3.1
 - Copilot