Functions Performance with Pass Rate

Pass Rate(%) vs Tasks

| Task | GPT-4o | GPT-4o-Mini | Gemini | Llama-3.1 | Copilot |
|------|--------|-------------|--------|-----------|---------|
| Remote Control | 69.6% | 52.2% | 13.0% | 0.0% | 39.1% |
| Fixed Plans | 81.5% | 43.9% | 24.2% | 0.0% | 31.8% |
| Auto Adapt | 57.7% | 17.9% | 6.4% | 0.0% | 0.0% |
| Energy Saving | 61.9% | 28.6% | 14.3% | 0.0% | 9.5% |