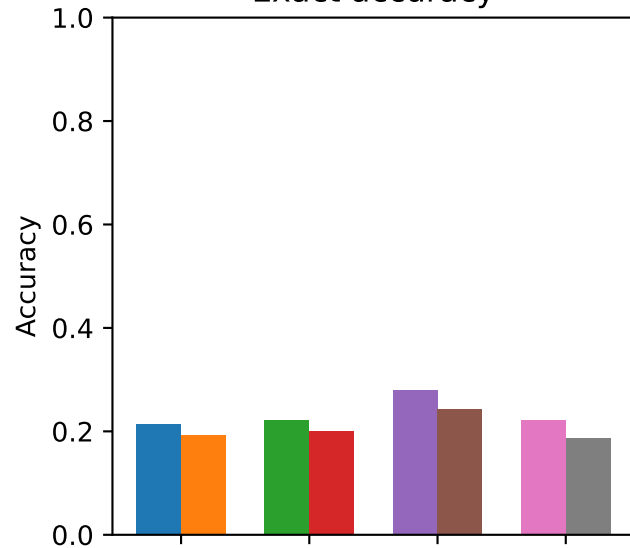
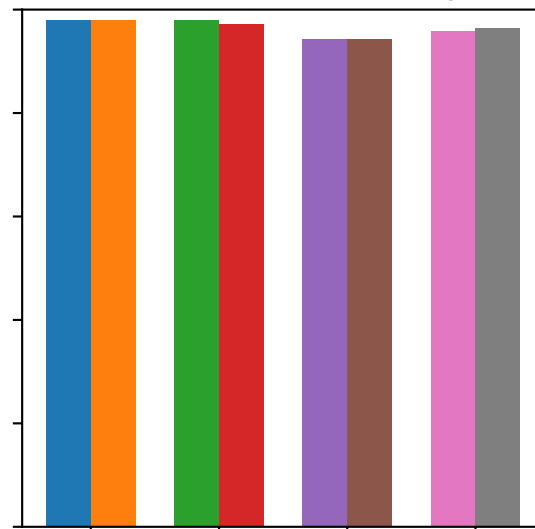


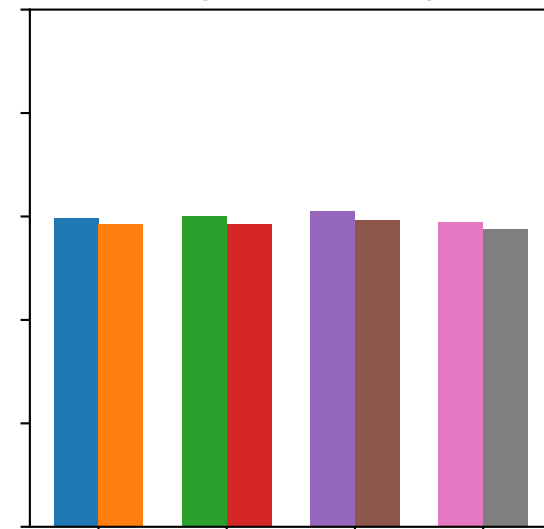
Exact accuracy



Conservative accuracy



Weighted accuracy



Max (Gemini 3 Flash, GPT-5-mini)

Gemini 3 Flash

GPT-5-mini

GPT-5.2

Max (Gemini 3 Flash, GPT-5-mini) (no search)
Max (Gemini 3 Flash, GPT-5-mini) (search)

Gemini 3 Flash (no search)
Gemini 3 Flash (search)

GPT-5-mini (no search)
GPT-5-mini (search)

GPT-5.2 (no search)
GPT-5.2 (search)