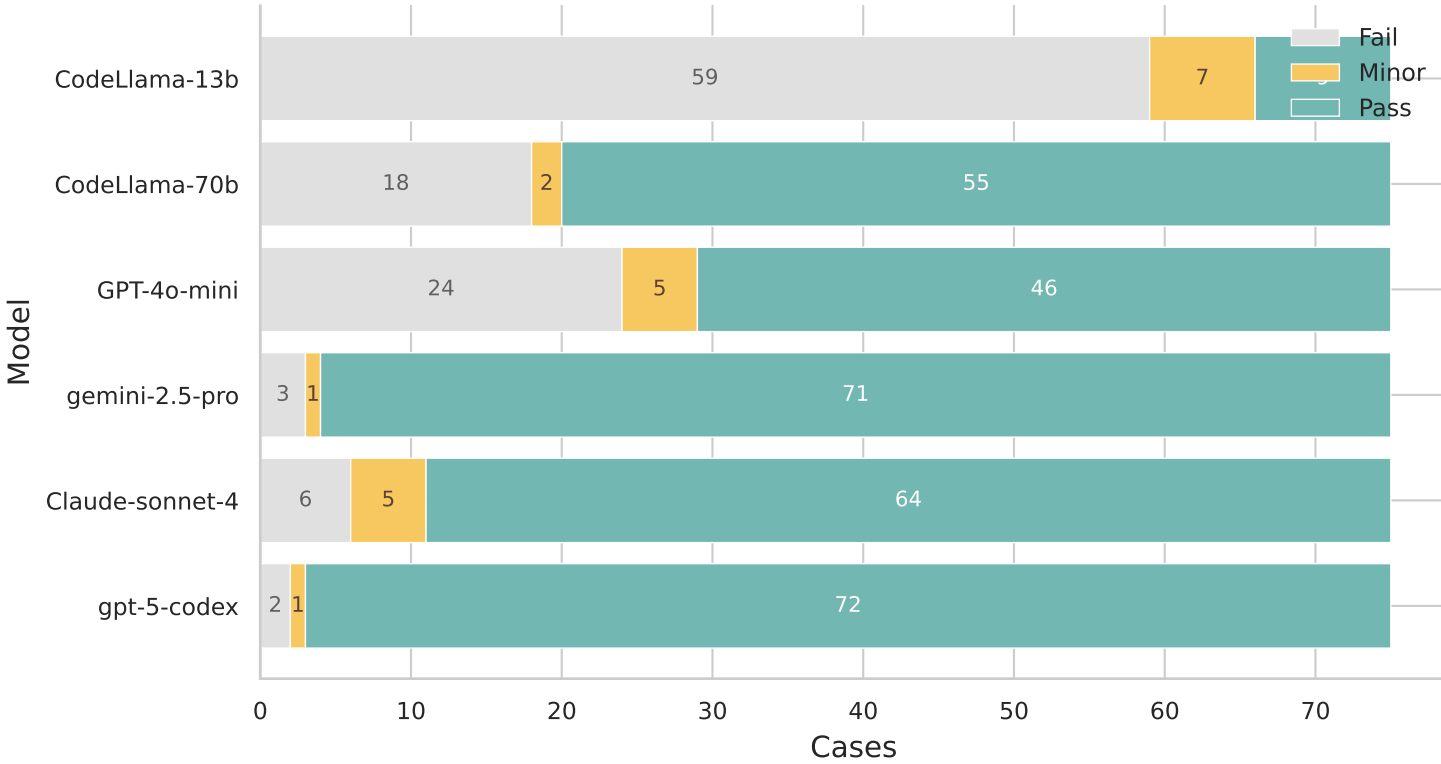
Fail
Minor
Pass
CodeLlama-13b
59
7
CodeLlama-70b
18
2
55
GPT-4o-mini
24
5
46
gemini-2.5-pro
3
1
71
Claude-sonnet-4
6
5
64
gpt-5-codex
2
1
72
Model
0
10
20
30
40
50
60
70
Cases