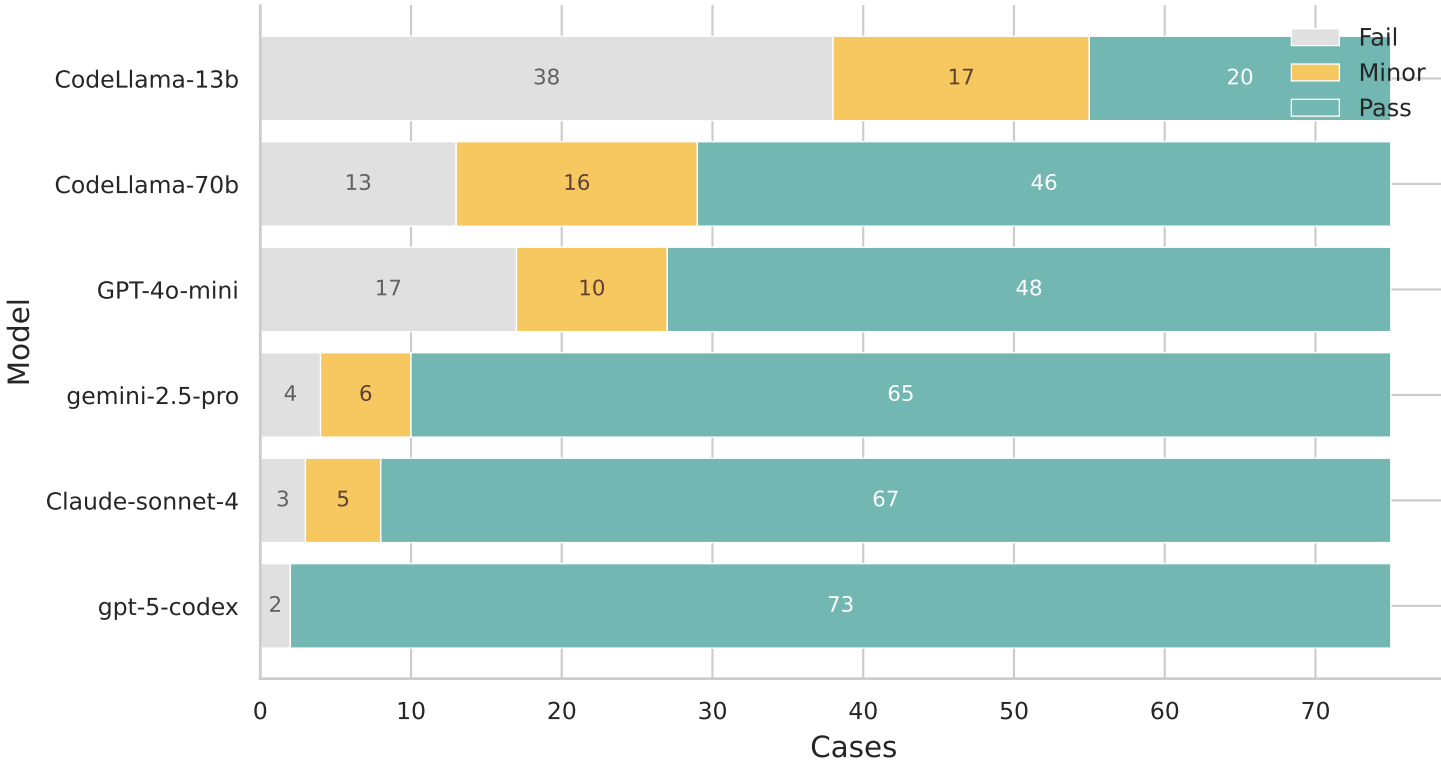
Fail
Minor
Pass
CodeLlama-13b
38
17
20
CodeLlama-70b
13
16
46
GPT-4o-mini
17
10
48
gemini-2.5-pro
4
6
65
Claude-sonnet-4
3
5
67
gpt-5-codex
2
73
Model
0
10
20
30
40
50
60
70
Cases