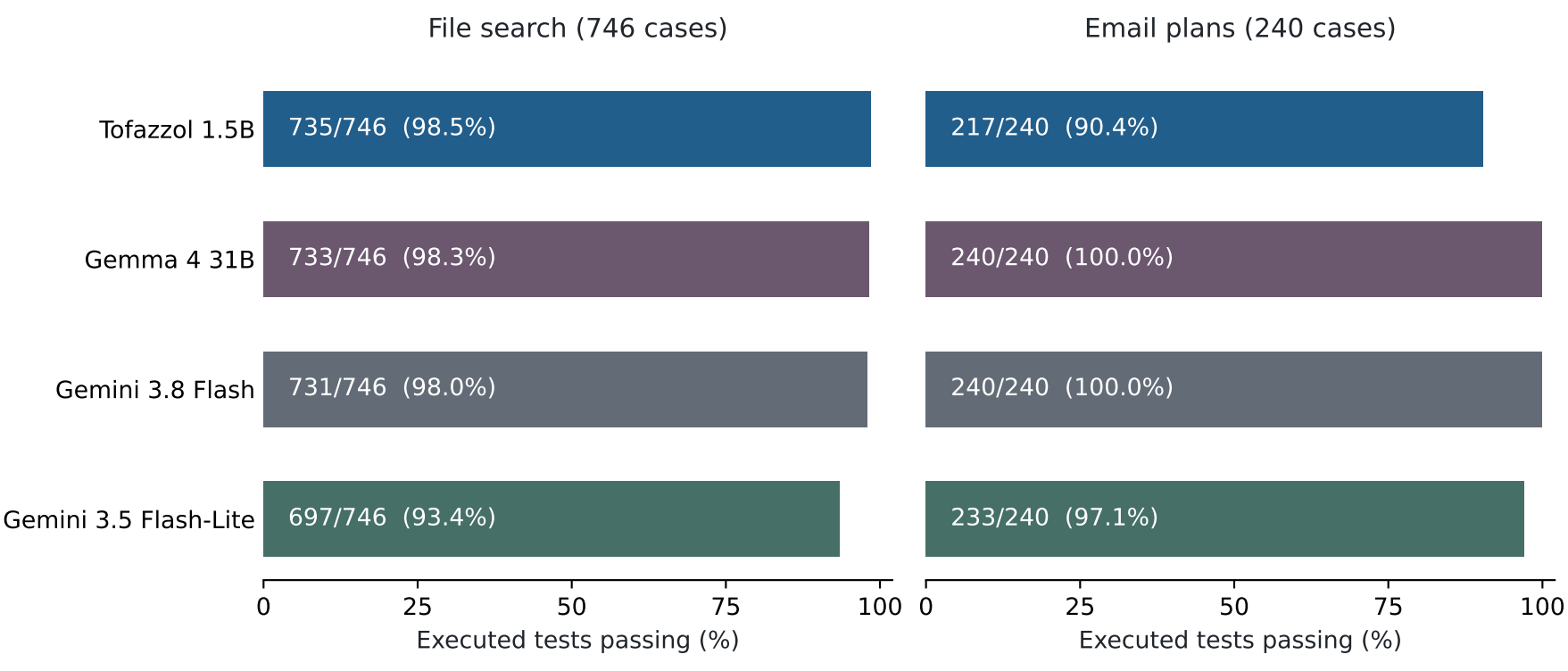

File search (746 cases)
Email plans (240 cases)
Tofazzol 1.5B
Gemma 4 31B
Gemini 3.8 Flash
Gemini 3.5 Flash-Lite
735/746 (98.5%)
733/746 (98.3%)
731/746 (98.0%)
697/746 (93.4%)
217/240 (90.4%)
240/240 (100.0%)
240/240 (100.0%)
233/240 (97.1%)
0
25
50
75
100
Executed tests passing (%)