Compiled successes by parameter count across LLMs (baseline filtered) 315 297 253 234 gpt-3.5 gpt-4o o1-mini (old) o3-mini deepseek claude3 sonnet 127 102 77 60 57 26

Parameter count (from untyped baseline)

successfully 5250 5250

Number of files typechecked