

# Large Language Monkeys (Benchmark = MiniF2F-MATH)

