# SWE-bench: percent solved