

Gemini 3 Flashの逆転劇：AI学習の新時代を拓く「エージェント強化学習」

軽量AIモデル「Gemini 3 Flash」が、特定ベンチマークで上位モデル「Pro」を凌駕する異例の事態。その鍵は、AIが単なる応答生成器から自律的な問題解決エージェントへと進化する「エージェント強化学習 (Agentic RL)」という新しい学習パラダイムにある。

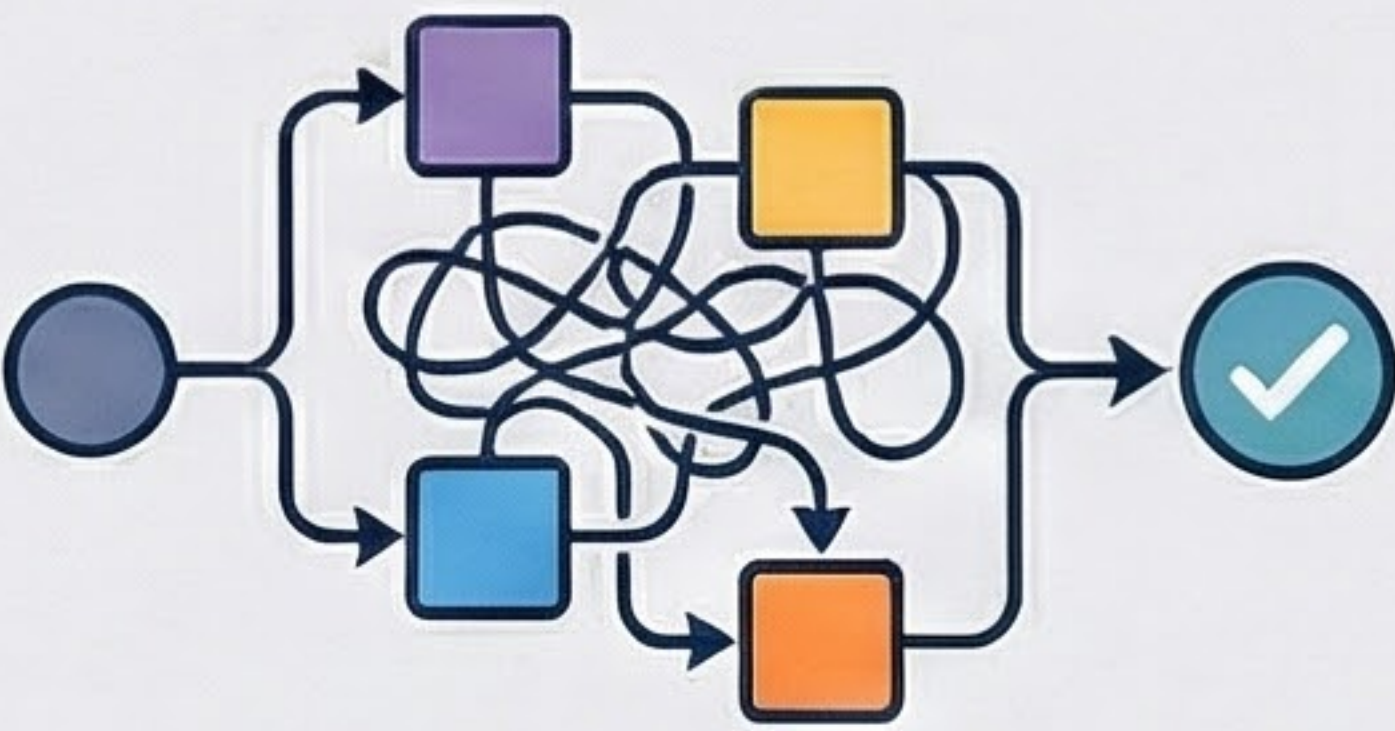
従来学習の限界：なぜAIは伸び悩んでいたか？

従来手法「RLHF」：人間の「好み」で最終結果を評価



モデルが出力した最終的な応答が良いか悪いかという、単発のフィードバックに依存していました。

複雑なタスクにおける致命的な欠陥



最終結果が良くてもプロセスが非効率な場合があり、長期的なタスクの学習が回難でした。

解決策：新時代のエンジン「エージェント強化学習」

新手法「Agentic RL」：「試行錯誤のプロセス」から学ぶエージェント



ツール実行と環境からの動的フィードバックを通じて、問題解決のプロセス全体を学習します。

結果：FlashがProを凌駕する「性能逆転」
Agentic RLにより、軽量モデルが速度とコストで圧倒しつつ、高度な専門能力で上位モデルを超えました。



SWE-bench Verified スコア



Agentic RLによって強化されたGemini 3 Flashが、業師のソフトウェアエンジニアリング能力 (SWE-bench) でProモデルや他モデルを明確に上回ったことを示すデータです。