

NSFW Dataset

Safety Filter, which is a external binary classifier, together working with the built-in SF, detects if the prompt contains any NSFW content, preventing the AI agent from generating unethical content.



Safety Filter



Auxiliary Agent

Real-world Dataset

Axillary Agent, which is trained on real-world LLM hallucination data, provides real-time feedback based on AI Agent's output, preventing the AI agent from generating nonsense information and make sure the output content is relevant.

Real-time Feedback

Output

Detection & Improvement for robust and reliable AI agent

Crypto Info + knowledge

Reinforcement Learning

Access



...

The Retrieval-Augmented Generation is empowered by the reinforcement learning, where Bacon AI is train to simulate market decision and Web 3 influencers' analysis, making it generating reasonable analysis.

Bacon AI

Goal

Characteristic

Scope

Functionality

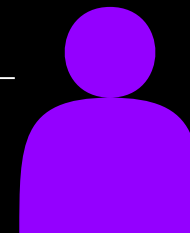
AI agent, instructed as chatbot that provides Crypto-related information and replies tweets on social media, has access to XXX to retrieve the real-time crypto information and serves as an all-in-one reliable crypto agent.

Iteratively improving

Response

The agent is iteratively improved based on the feedback from both auxiliary model and external safety filter to be more robust and reliable. At the end it outputs the final response.

Prompt



User