Inter-adversary Prompting

update strategy

Detection Explanations

Generator — LLM — Fake Strategies

Fake News — FAKE

Detector — LLM — Detection Strategies

if ✓ Correct

else ✗ Incorrect

Fake Explanations — update strategy

Collected Real News

Self-reflection Prompting

News

Detector — LLM — Detection Strategies

if ✓ Correct

else ✗ Incorrect

update strategy

Reflection

Reflector — LLM

Detection Explanations

Explainable Fake News Detection

News

Detector — LLM — Detection Strategies

News: Detailed photos of Xiang Liu's tendon surgery exposed. Show sympathy and blessings!
Prediction: Fake
Explanations: Real surgery generally won't be exposed...