Analytics Vidhya

Verify-and-Edit Prompting

Instructor

Bhaskarjit Sarmah

Vice President, BlackRock
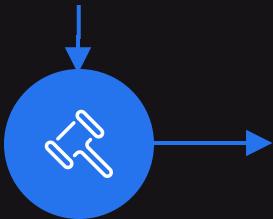
Definition
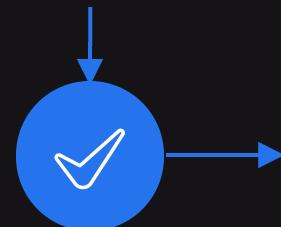
- **Verify-and-Edit Prompting** is a method that enhances answer accuracy by using external sources to verify and refine responses.

- This method goes beyond Self-consistency by incorporating validation from resources like Wikipedia or Google.

How Does it Work?
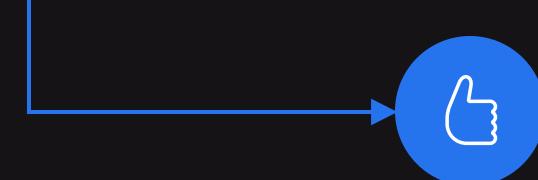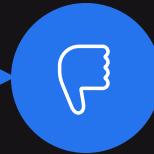
LLM assesses its responses for certainty before seeking external information

If initial certainty is lacking, LLM queries for more information

Utilize new data to correct and improve its initial answers

Analytics Vidhya

Verify-and-Edit Prompting: Example

Question

Of all the teams John Nyskohus played for, which team was known as the 'Black and Whites'?

Standard

Newcastle United ❌

Chain-of-thought

First, John Nyskohua played for the Norwegian football team Odd Grenland. Second, Odd Greenland is known as 'the Black and Whites'. The answer is Odd Greenland. ❌

Self-consistency

Verify
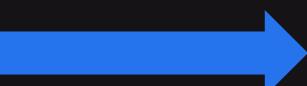
What team did John Nyskohus play for?
What team is known as 'the Black and Whites?'

External Knowledge Retrieval

John Nyskohus ..... is an Australian former soccer player who played club football for USC Lion .... and Adelaide City in the National Soccer League...
Adelaide City Football Club is an Australian football club based in Adelaide, south Australia. They are also known as 'The Zebras" and 'the Black and Whites'.

Edit Rationales

First, John Nyskohus played for Adelaide City in the National Soccer League. Second, Adelaide City Football Club is known as 'the Black and Whites.'

The answer is Adelaide City Football Cub. ✓

Pros: Verify-and-Editing Prompting

Increases LLM accuracy by fact-checking initial guesses.

Clarifies the LLM's thought process for greater transparency.

Uses real-world data, enhancing trust and applicability.

Outperforms other methods in complex problem-solving.

Cons: Verify-and-Edit Prompting

Most effective for complex questions; less impact on simpler, commonsense queries.

Depends on a reliable verification system to improve answer quality.

Risks introducing errors or irrelevant data during the edit process.

Thank You