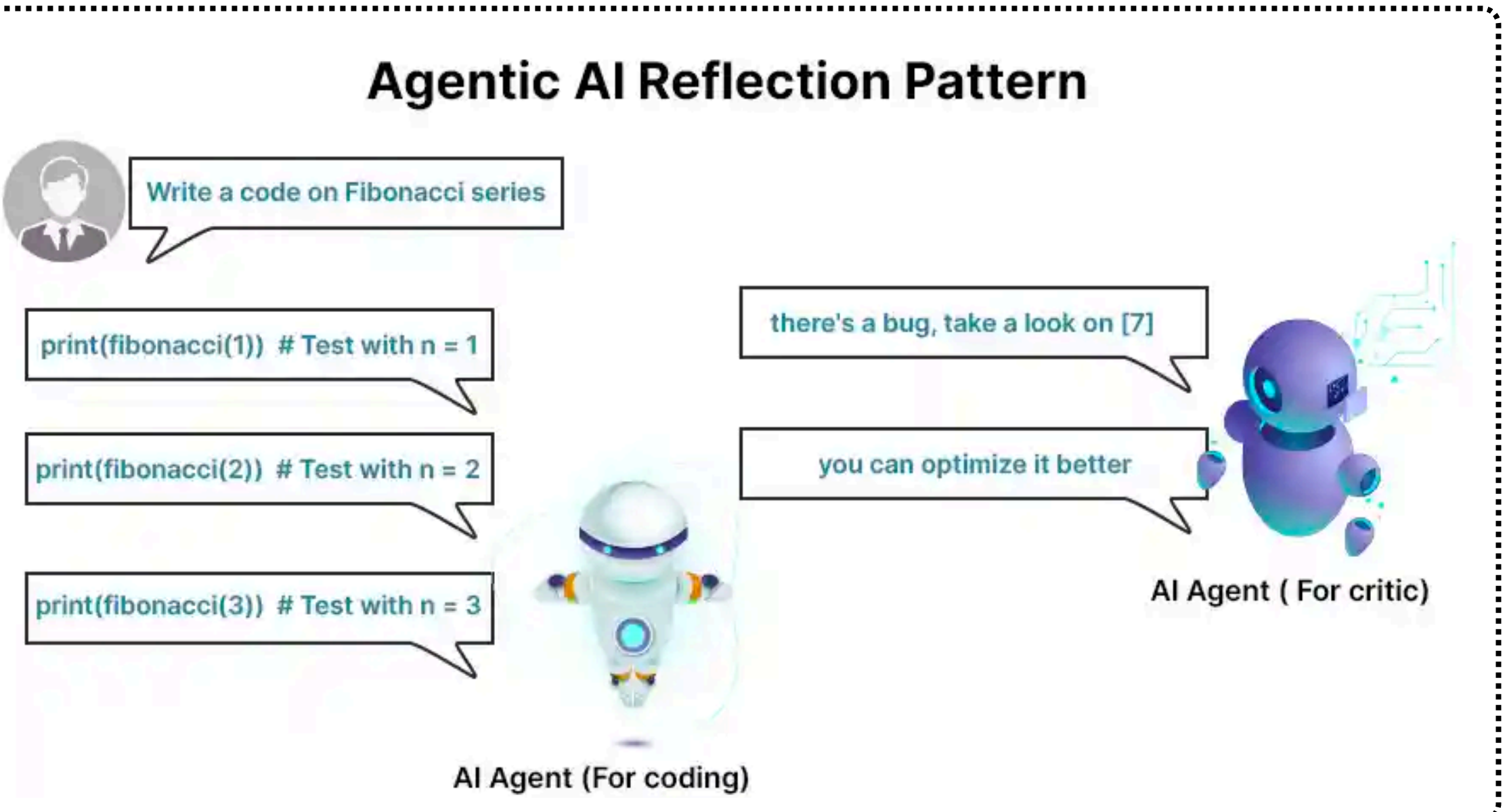
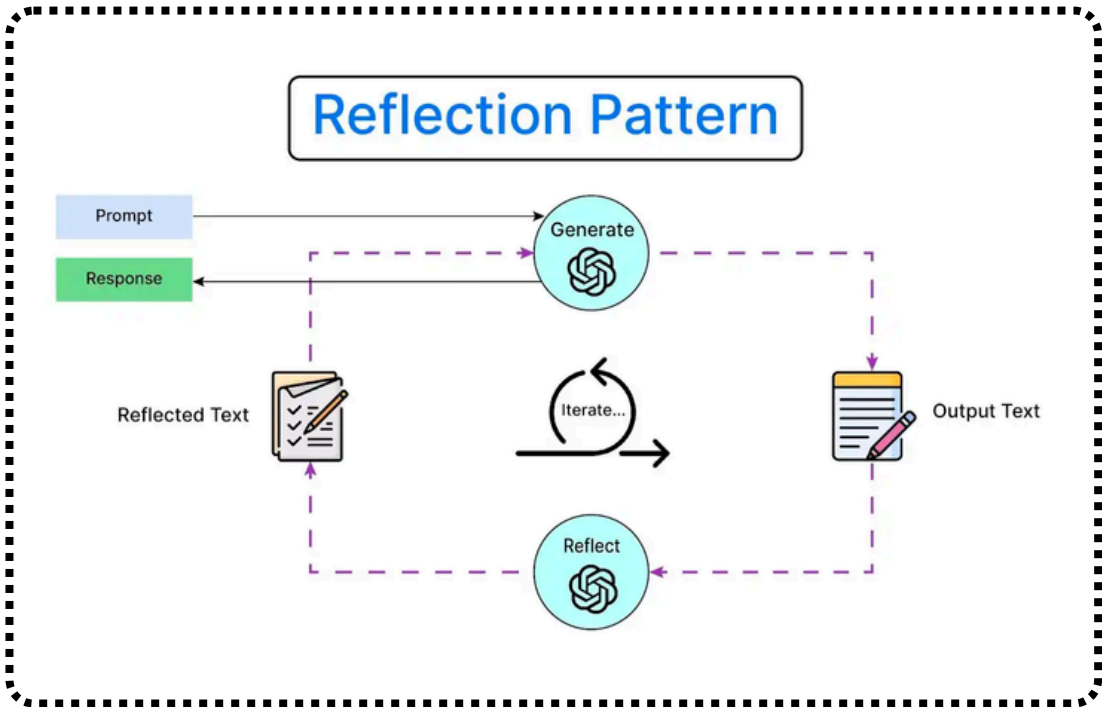
Analytics Vidhya

Day 10 of Mastering AI Agents

Agentic AI Reflection Pattern






```
reflection_chat_history = [
   {
   "role": "system",
   "content": "You are Nitika Sharma, an experienced Python coder. With this experience in Python
generate critique and recommendations for user output on the given prompt",
   }
]
reflection_chat_history.append(
   {
       "role": "user",
       "content": fibonacci_code
   }
)
critique = client.chat.completions.create(
   messages=reflection_chat_history,
   model="llama3-70b-8192"
).choices[0].message.content
display_markdown(critique, raw=True)
```

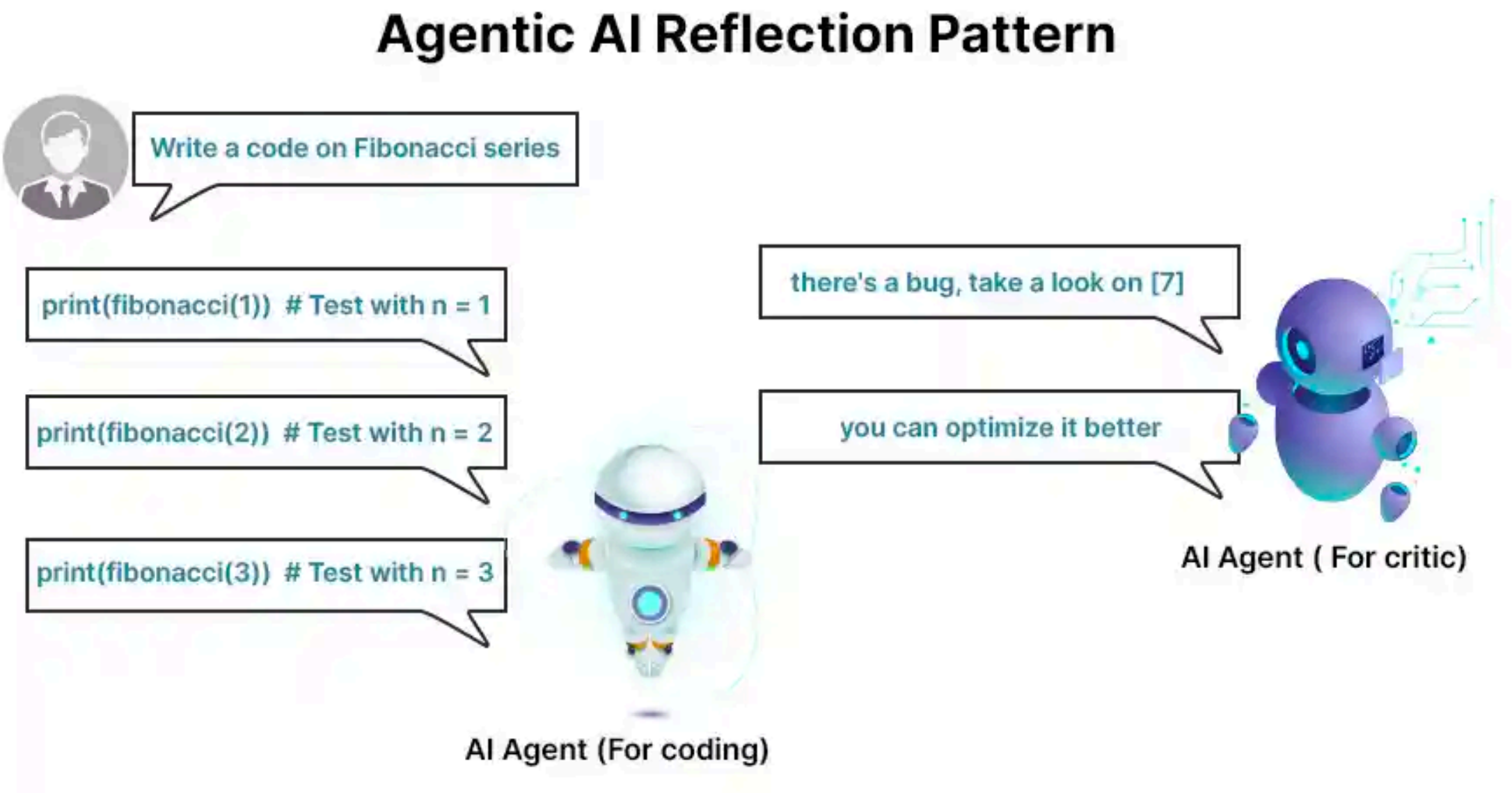
What is the Reflection Pattern?

The Reflection Pattern is an agentic AI design pattern applied to AI models, where the model generates an initial response to a prompt, evaluates this output for quality and correctness, and then refines the content based on its own feedback. The model essentially plays the dual roles of creator and critic. The process involves several iterations where the AI alternates between these two roles until the output meets a certain level of quality or a predefined stopping criterion.

It evaluates its own work, checks for errors, inconsistencies, or areas where the output could be enhanced, and then makes revisions. This cycle of generation and self-assessment allows the AI to refine its responses iteratively, leading to much more accurate and useful results over time.

This pattern is especially valuable for large language models (LLMs) because language can be complex and nuanced. By reflecting on its own outputs, the AI can catch mistakes, clarify ambiguous phrases, and ensure that its responses better align with the intended meaning or task requirements. Just like our course developer refining lessons to improve learning outcomes, the Reflection Pattern enables AI systems to improve the quality of their generated content continuously.

Why Use the Reflection Pattern?



The reflection pattern is effective because it allows for incremental improvement through iterative feedback. By repeatedly reflecting on the output, identifying areas for improvement, and refining the text, you can achieve a higher-quality result than would be possible with a single generation step.

Imagine using this pattern when writing a research summary.

- Prompt: “Summarize the key points of this research paper on climate change.”
- Generate: The AI provides a brief summary.
- Reflect: You notice that some important aspects of the paper, such as the implications of the findings, are missing.
- Reflected Text: You update the summary to include these details and refine the language for clarity.
- Iterate: You repeat the process until the summary accurately captures all the critical points.

This approach encourages continuous refinement and is particularly useful in complex tasks such as content creation, editing, or debugging code.

Let’s discuss the key components of the the Reflection Pattern in the next slide------->

Key Components

The Reflection Pattern consists of three main components:

1. Generation Step

The process begins when a user provides an initial prompt, which could be a request to generate text, write code, or solve a complex problem. For example, a prompt might ask the AI to generate an essay on a historical figure or to implement an algorithm in a specific programming language.

- Zero-Shot Prompting: The first generation is often done in a zero-shot style, where the AI generates a response without previous examples or iterations.

- Initial Output: The output produced is considered a first draft. While it may be relevant and coherent, it may still contain errors or lack the necessary detail.

The goal of the generation step is to produce a candidate output that can be further evaluated and refined in subsequent steps.

2. Reflection Step

The reflection step is a critical phase where the AI model reviews its own generated content. This step involves:

- **Self-Critique**: The model critiques its own work, identifying areas for improvement, such as factual errors, stylistic issues, or logical inconsistencies.

- **Feedback Generation**: The AI generates specific feedback, which can include suggestions for restructuring content, adding details, or correcting mistakes.

- **Evaluation Criteria**: The critique may be based on predefined criteria such as grammatical accuracy, coherence, relevance to the prompt, or adherence to specific formatting guidelines.

The reflection process can involve mimicking the style of a subject matter expert to provide more in-depth feedback. For instance, the AI might adopt the persona of a software engineer to review a piece of code or act as a historian critiquing an essay.

3. Iteration and Refinement

In this phase, the feedback generated during the reflection step is used to guide the next generation of output. The AI incorporates the suggested changes and improvements into a new version of the content. This cycle repeats multiple times, with each iteration bringing the output closer to the desired quality.

- **Adaptive Learning**: Through this iterative process, the AI learns to recognize patterns in its own mistakes and refines its understanding of the task requirements.

- **Multiple Iterations**: The process can be repeated for a fixed number of steps (e.g., 10 iterations) or until a specific stopping condition is met, such as achieving a certain level of content quality or encountering a “stop” keyword.

Advanced AI Agents Artificial Intelligence Large Language Models LLMs

What is Agentic AI Reflection Pattern?

Agentic AI Reflection Pattern improves AI systems through iterative self-assessment, refining outputs for enhanced accuracy & more.

Pankaj Singh 02 Jan, 2025