

NAME : JESINTHA R  
REG NO :212222060096

## **Evaluation of Prompting Tools Across Diverse AI Platforms**

### **Aim:**

To assess the effectiveness, usability, and response quality of various prompting tools across popular AI platforms—ChatGPT, Claude, Bard, Cohere Command, and Meta—in a specific use case (e.g., summarizing text or answering technical questions).

### **Procedure:**

- Step 1: Define the Use Case

#### Objective:

Identify a common, complex task to evaluate the AI platforms consistently.

#### Use Case Selected:

-  Answering Complex Technical Questions

#### Sample Task:

“Explain convolutional neural networks (CNNs) and their applications in image processing.”

This task requires understanding of deep learning, image processing concepts, and the ability to explain them clearly.

---

- Step 2: Design Standardized Prompts

To ensure fair and consistent evaluation, standardized prompts are created for the use case.

#### Examples of Prompts for Different Use Cases:

- Summarization Prompt:

“Summarize the following document about machine learning techniques.”

- Technical Q&A Prompt (Chosen):

“Explain convolutional neural networks (CNNs) and their applications in image processing.”

- Creative Writing Prompt:

“Generate a short story about an AI robot in a futuristic city.”

These prompts will be used across all selected AI platforms without modification.

---

### Step 3: Input Prompts into Each Platform

Platforms for Evaluation:

1. OpenAI ChatGPT
2. Anthropic Claude
3. Google Bard / Gemini
4. Cohere Command R+
5. Meta AI (LLaMA or similar)

Procedure:

- Input the same prompt into each platform.
  - Record the full response generated.
  - Capture metadata such as response time and version used (if applicable).
- 

### Step 4: Evaluate Response Quality

Each response is assessed using four key criteria:

Criteria      Description

Accuracy      Correctness of the information and adherence to the prompt.

Clarity      Logical structure, readability, and ease of understanding.

Depth      Thoroughness, detailed explanation, and inclusion of relevant examples.

Creativity      (If applicable) Originality and engagement, especially for creative tasks.

Rating Scale:

Each criterion is rated from 1 (Poor) to 5 (Excellent).

### Example Evaluation for CNN Prompt:

Platform	Accuracy	Clarity	Depth	Avg Score
ChatGPT	5	5	5	5.0
Claude	5	5	4	4.7
Bard (Gemini)	4	4	4	4.0
Cohere	4	3	3	3.3
Meta AI	4	3	3	3.3

---

## Step 5: Measure User Experience

In addition to output quality, the user interaction with the platform is evaluated using the following metrics:

Metric	Description
Response Speed	Time taken for the AI to generate a full response.
Interface Usability	Intuitiveness and ease of navigating the platform.
Prompt Customization	Ability to tweak prompts or get clarification interactively.
Interaction Features	Support for follow-up questions, session memory, and conversational context.

### Example UX Ratings:

Platform	Speed	Usability	Customization	Interaction	Avg UX Score
ChatGPT	5	5	5	5	5.0
Claude	5	4	4	4	4.3
Bard (Gemini)	4	4	4	4	4.0
Cohere	4	3	3	3	3.3
Meta AI	4	3	2	3	3.0

---

## Step 6: Compare Consistency Across Queries

To measure how well platforms maintain performance:

1. Run multiple prompts under the same technical domain.
2. Observe any variations in:
  - o Information quality
  - o Style and tone
  - o Interface reliability

Additional Prompts for Consistency Testing:

- “What is transfer learning in CNNs?”
- “How do CNNs compare with RNNs for sequential data?”
- “Describe the role of pooling layers in CNNs.”

 Consistent Platforms: ChatGPT and Claude

## ⚠️ Variable Platforms: Bard, Cohere, Meta AI

---

### 📈 Final Comparison Summary

Platform	Quality Score	UX Score	Consistency	Verdict
----------	---------------	----------	-------------	---------

ChatGPT	5.0	5.0	High	✅ Best Overall for Tech Tasks
Claude	4.7	4.3	High	✅ Strong Alternative
Bard	4.0	4.0	Moderate	⚠️ Good for General Use
Cohere	3.3	3.3	Low	⚠️ Still Developing
Meta AI	3.3	3.0	Low	⚠️ Basic Use Only

### 📋 Additional Notes

- Use ChatGPT or Claude for academic, research, or engineering-level queries.
- Bard (Gemini) offers good language fluency but may lack depth in technical topics.
- Cohere and Meta AI are improving rapidly but currently limited for complex reasoning.
- Re-evaluate regularly — these models are updated frequently

## Outcomes:

**Prompt: “Explain convolutional neural networks (CNNs) and their applications in image processing.”**

The screenshot shows a ChatGPT interface. At the top, there are icons for a profile picture, a message, and a dropdown menu labeled "ChatGPT". To the right of the menu is a "Memory Full" indicator with a circular arrow icon. The main area has a light gray background. A user input box contains the prompt: "Explain convolutional neural networks (CNNs) and their applications in image processing." Below the input box, the AI's response is displayed in a card-like format. The response starts with a section titled "Convolutional Neural Networks (CNNs):" which defines CNNs as a deep learning model for image and video data. It then moves to a section titled "How CNNs Work:" which includes a numbered list under "1. Convolution Layer:". The list describes how convolution layers apply filters across an image to detect patterns like edges, textures, and shapes. It also explains that each filter performs a convolution operation to highlight specific features. The final point in the list states that this layer captures local features by focusing on small regions of the image. There is a downward-pointing arrow icon next to the list. Below this, another section titled "2. Activation Function (usually ReLU):" is partially visible. At the bottom of the screen, there is a footer bar with a "Message ChatGPT" button and an upward-pointing arrow icon. The footer also includes a note: "ChatGPT can make mistakes. Check important info."

Claude

Claude does not have the ability to run the code it generates yet.

PA Explain convolutional neural networks (CNNs) and their applications in image processing

I'll explain Convolutional Neural Networks (CNNs) and their role in image processing in a clear, structured way.

Understanding Convolutional Neural Networks (CNNs)  
Click to open document

CNN Architecture Diagram  
Click to open image

I've provided a comprehensive explanation of CNNs along with a visual representation of their architecture. Would you like me to elaborate on any specific aspect of CNNs or provide more details about a particular application area?

Claude can make mistakes. Please double-check responses.

Reply to Claude...

Claude 3.5 Sonnet (New)

CNN Architecture Diagram

Last edited just now

Preview Code X

Gemini

See the latest updates to the Gemini Apps Privacy Hub

P Explain convolutional neural networks (CNNs) and their applications in image processing.

Show drafts v 🔍

Convolutional Neural Networks (CNNs): A Visual Guide

Convolutional Neural Networks (CNNs) are a type of artificial neural network specifically designed to process visual imagery. They're inspired by the structure and function of the human visual cortex. CNNs are the backbone of many image recognition and processing tasks, from facial recognition to medical image analysis.

How CNNs Work

A typical CNN consists of several layers:

1. Convolutional Layer:
  - The core of a CNN.
  - Applies filters (kernels) to the input image, sliding them across the image to extract features like edges, corners, and textures.

Ask Gemini

Gemini can make mistakes, so double-check it

cohere

Chat below to try Command R+, now with 10 supported languages!

CHAT DASHBOARD PLAYGROUND DOCS COMMUNITY

Chats

PINNED (0)

MOST RECENT (1)

Convolutional Neural Networks (CNNs) and Their Applications in Image Processing

P Explain convolutional neural networks (CNNs) and their applications in image processing.

Convolutional Neural Networks (CNNs) are a class of artificial neural networks that have revolutionized the field of image processing and computer vision. They are designed to automatically learn and extract meaningful features from images, making them highly effective for various tasks such as image classification, object detection, image segmentation, and more. Here's an explanation of CNNs and their applications:

1. Convolutional Layers:
  - CNNs get their name from the convolutional layers, which are the core building blocks of these networks. Convolutional layers apply a mathematical operation called convolution to the input image.
  - In a convolution, a small set of weights called a "kernel" or "filter" is slid across the image, performing element-wise multiplication and summation with the image pixels. This process captures local patterns and features within the image.
  - The output of the convolutional layer is a feature map that highlights specific patterns, such as edges, textures, or shapes, present in the input image.
2. Feature Extraction:

Message... ➔

Message... ➔

The screenshot shows the Meta AI web interface. At the top left is the Meta AI logo and a "New conversation" button. The main area has a dark background with white text. A user message at the top right reads: "Explain convolutional neural networks (CNNs) and their applications in image processing." Below this, the AI's response is: "Convolutional Neural Networks (CNNs) are a type of deep learning algorithm inspired by the structure and function of the human visual cortex. They are primarily designed for image and video processing tasks." To the right of the text are four small circular icons with symbols: a grid, a thumbs up, a thumbs down, and a reply arrow. On the left side of the main area, there's a sidebar with a "AI Demos" section containing a "Log in" button and links to "Privacy · Terms · AI Terms · Cookies". At the bottom of the sidebar, it says "Meta © 2024". The bottom right of the main area has a "Ask Meta AI anything..." input field and a note: "Messages are generated by AI and may be inaccurate or inappropriate. Learn more".

## Conclusion:

The comparison reveals strengths and weaknesses of each AI platform in the chosen use case. While some platforms excel in **accuracy and depth** for technical prompts (like ChatGPT and Claude), others may outperform in **creativity** for storytelling (e.g., Meta).

Factors like **interface usability** and **response customization options** also influence the user experience, with tools like Claude and Meta offering flexible prompt modifications, which can be beneficial for refining responses.

This evaluation provides a comprehensive understanding of each platform's capabilities and ideal use cases, guiding users in selecting the best tool based on specific needs, whether technical, creative, or interactive.