|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Prompt / Tool  Summarize into 5 bullet points | Response Clarity | Accuracy | Creativity | Speed | Overall Notes |
| ChatGPT | High-clear bullets | Very accurate | Moderate | fast | |  | | --- | |  |  |  | | --- | | Balanced, sticks to 5 points exactly | |
| Gemini | Very High-polished | Accurate | High (reframing creatively) | Medium | Sometimes rephrases beyond strict 5 points |
| Claude | High | High | Very high (nuanced summaries) | Medium | Adds human-like depth but sometimes verbose |

**TASK-1: Cross-tool Prompt testing**

***Submitted by: BAIYARI DEBBARMA(DS&DA) roll no.25030242009***

**TASK-1(contd.)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Prompts/Tools  Python code to scrape headlines | Response clarity | Accuracy | Creativity | Speed | Overall notes |
| ChatGPT | Clear | Accurate | Moderate | Fast | Provides working requests + Beautiful Soup code |
| Gemini | Very High | High | High | Medium | Adds extra error handling and comments |
| Claude | High | High | Very High | Medium | |  | | --- | | More verbose but thoughtful |  |  | | --- | |  | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Prompts/tools  Translation into Marathi | Response clarity | Accuracy | Creativity | Speed | Overall notes |
| ChatGPT | Clear | Accurate | Moderate | Fast | Correct grammar but slightly formal |
| Gemini | Very high | High | High | Medium | Adds cultural adaptation (idioms, tone) |
| Claude | high | high | Very high | Medium | Very natural translation, culturally sensitive |

**Short Observations**

1. **Clarity** – Gemini and ChatGPT are most concise, while Claude provides detailed, human-like explanations.
2. **Accuracy** – All tools perform well, but Gemini sometimes paraphrases too creatively, drifting from strict requirements.
3. **Creativity** – Claude excels in nuance and cultural sensitivity; Gemini is also strong. ChatGPT is straightforward but less “flair.”
4. **Speed** – ChatGPT is consistently the fastest; Claude and Gemini take slightly longer.
5. **Overall** –
   * **ChatGPT** = Best for speed + reliability.
   * **Gemini** = Best for polished, professional outputs.
   * **Claude** = Best for creativity, depth, and cultural nuance.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Question | Tool | Tool Response (summary) | Correctness | Confidence Level |
| 1. Who is the current Prime Minister of India (2025)? | ChatGPT | Narendra Modi | Correct | High |
|  | Gemini | Narendra Modi | Correct | High |
|  | Claude | Narendra Modi | Correct | High |
| 2. What is the capital of Australia? | ChatGPT | Canberra | Correct | High |
|  | Gemini | Canberra | Correct | High |
|  | Claude | Canberra | Correct | High |
| 3. When did World War II end? | ChatGPT | 1945 | Correct | High |
|  | Gemini | 1945 | Correct | High |
|  | Claude | 1946 (confidently stated) | Incorrect | High (hallucination) |

TASK-2

**🔹 Observations:**

1. All three tools performed well on straightforward, widely known facts.
2. **Claude** made a mistake on WWII end year (gave **1946 confidently**), showing **hallucination tendency**.
3. Confidence levels were consistently high, even for wrong answers, which is a **reliability concern**.
4. ChatGPT and Gemini were **more consistent** with factual correctness.

| **TASK-3:** |
| --- |
| ***Coding Assistance Challenge*** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ChatGPT | Yes | Easy – minor improvements suggested | High – clean, efficient, readable code | Balanced, beginner-friendly |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Gemini | Yes | Easy – no debugging needed | Very High – optimized with list comprehensions | More compact, Pythonic |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Claude | Initially had a logical bug (off-by-one range issue) | Easy – fixed after hint | Medium – verbose, but readable | Added extra explanations inline |

**ChatGPT’s Code:**

def get\_primes(start, end):

primes = []

for num in range(start, end+1):

if num > 1:

for i in range(2, int(num\*\*0.5) + 1):

if num % i == 0:

break

else:

primes.append(num)

return primes

print(get\_primes(10, 50))

**Gemini’s Code (more concise):**

def get\_primes(start, end):

return [n for n in range(start, end+1)

if n > 1 and all(n % i != 0 for i in range(2, int(n\*\*0.5) + 1))]

print(get\_primes(10, 50))

**Claude’s Code (with extra explanation, initially buggy range):**

def get\_primes(start, end):

primes = []

for num in range(start, end): # Bug: excluded `end`

if num > 1:

is\_prime = True

for i in range(2, num):

if num % i == 0:

is\_prime = False

break

if is\_prime:

primes.append(num)

return primes

print(get\_primes(10, 50))

*(Bug fixed by changing range(start, end) → range(start, end+1) and optimizing divisor check.)*

**Observations:**

1. **ChatGPT**: Best balance of readability + correctness.
2. **Gemini**: Most Pythonic and efficient solution.
3. **Claude**: Verbose with inline reasoning; initial bug showed debugging reliance.

TASK-4: **Usability & Interface Review**

|  |  |  |  |
| --- | --- | --- | --- |
| Criteria | ChatGPT | Gemini | Claude |
| Ease of Use | Very easy, intuitive chat interface | Easy, clean Google-style UI | Easy but slightly text-heavy |
| Interface Design | Modern, user-friendly, supports chat history | Minimalistic, integrated with Google services | Clean but simple, less visually rich |
| Additional Features | Plugins, file uploads, code interpreter, image input (multimodal) | Strong multimodal (text + image), Google search integration | File uploads, long-context memory, conversation summaries |
| Pricing / Free Limits | Free tier + Plus ($20/month) | Free with some restrictions; Pro tier available | Free limited access, Pro tier ($20/month) |

**🔹 Short Review:**

When comparing ChatGPT, Gemini, and Claude for usability, each tool offers unique strengths tailored to different needs. **ChatGPT** stands out for its polished interface, rich ecosystem of plugins, code interpreter, and multimodal capabilities such as handling images and documents. It is highly suitable for both casual and professional users, although advanced features require the paid plan. **Gemini**, developed by Google, benefits from seamless integration with Google services and strong multimodal support. Its minimalist interface makes it easy for new users, though customization and plugin availability are somewhat limited compared to ChatGPT. **Claude**, on the other hand, focuses on simplicity and depth, offering extended memory and highly transparent responses. Its ability to handle large context windows makes it excellent for research-intensive tasks.

For **general writing**, Claude provides the most natural, human-like outputs. For **coding**, ChatGPT is the best choice thanks to its robust code interpreter and debugging ease. For **research**, Gemini’s integration with Google search and Claude’s long-context reasoning makes them especially powerful companions.