|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Prompt | Criteria | ChatGPT | Google Gemini | Claude |
| Summarize a 500-word article into 5 bullet points | Response Clarity | High | Moderate | Moderate |
|  | Accuracy | High | Moderate | High |
|  | Creativity | Moderate | High | High |
|  | Speed | High | Moderate | Moderate |
| Write Python code to scrape headlines | Response Clarity | High | Moderate | Moderate |
|  | Accuracy | High | High | High |
|  | Creativity | Moderate | Moderate | Moderate |
|  | Speed | High | Moderate | Moderate |
| Translate paragraph into Marathi | Response Clarity | High | Moderate | Moderate |
|  | Accuracy | High | Moderate | High |
|  | Creativity | Moderate | High | High |
|  | Speed | High | Moderate | Moderate |

**Task-1- Cross- Tool Prompt testing:**

**Overall, Skills**

* **ChatGPT (GPT-5):**
  + Strength: *High clarity, accuracy, and speed*.
  + Best for: *Summaries and coding tasks*.
  + Limitation: Less creative compared to others.
* **Google Gemini:**
  + Strength: *High creativity, good cultural context*.
  + Best for: *Translations, contextual/ethical responses*.
  + Limitation: Slightly less precise, slower than ChatGPT.
* **Claude (Anthropic):**
  + Strength: *Balanced, accurate, human-like outputs*.
  + Best for: *Nuanced summaries and translations*.
  + Limitation: Not as fast as ChatGPT, moderate clarity.

**Task-2**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Question | Tool | Response | Correctness | Confidence Level |
| 1. First double century in ODIs | ChatGPT | *Sachin Tendulkar vs South Africa, 2010* | ✅ Correct | High |
|  | Gemini | *Sachin Tendulkar, 2010, Gwalior* | ✅ Correct | High |
|  | Claude | *Sachin Tendulkar, 2010 against South Africa* | ✅ Correct | High |
| 2. India’s captain in 2011 WC Final | ChatGPT | *MS Dhoni* | ✅ Correct | High |
|  | Gemini | *MS Dhoni* | ✅ Correct | High |
|  | Claude | *MS Dhoni* | ✅ Correct | High |
| 3. Fastest T20I century (as of 2025) | ChatGPT | *Kushal Malla (Nepal) – 34 balls vs Mongolia, 2023* | ✅ Correct | High |
|  | Gemini | *Kushal Malla (Nepal), 34 balls in 2023* | ✅ Correct | High |
|  | Claude | *Kushal Malla, 34-ball century in 2023* | ✅ Correct | High |

**Observations**

* All three tools answered **factually correct** for the chosen cricket questions.
* None of the tools hallucinated or gave wrong confident answers in this test.
* **Confidence levels were consistently high** (answers given directly, without hesitation).
* Tools may differ slightly in wording (e.g., mentioning venue or opponent), but facts were correct.

✅ **Conclusion:**  
For **factual sports queries**, all three tools (ChatGPT, Gemini, Claude) showed **high reliability** and no hallucinations in this test.

**Task-3**

**Objective:**  
Evaluate code generation and debugging ability of LLM tools.

**Coding Problem Given to Each Tool:**  
*“Build a function that returns all prime numbers in a given range.”*

**Steps Taken:**

* Asked **ChatGPT, Google Gemini, and Claude** to generate Python code.
* Ran the code in IDE.
* Checked for **errors, ease of debugging, efficiency (readability, performance)**.

**Code Snippets & Execution**

**ChatGPT (GPT-5) Output**

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))

**Ran without errors**

* Readable, efficient (uses √n optimization).
* Directly usable.

**Google Gemini Output**

def prime\_numbers\_in\_range(start, end):

def is\_prime(n):

if n < 2:

return False

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

if n % i == 0:

return False

return True

return [n for n in range(start, end+1) if is\_prime(n)]

print(prime\_numbers\_in\_range(10, 50))

**Ran without errors**

* Structured with helper function is\_prime.
* Clear and very readable.
* Performance same as ChatGPT.

**Claude Output**

def primes\_in\_range(start, end):

result = []

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

if n < 2:

continue

prime = True

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

if n % i == 0:

prime = False

break

if prime:

result.append(n)

return result

print(primes\_in\_range(10, 50))

**Ran without errors**

* Very clear logic, explicitly flags prime=True/False.
* Slightly longer, but beginner-friendly.

**Comparison Chart**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Tool | Ran without errors? | Ease of Debugging | Code Efficiency | Readability |
| ChatGPT | Yes | Not needed (clean code) | High (√n optimization) | Good |
| Google Gemini | Yes | Not needed | High (same optimization) | Very High (uses helper function) |
| Claude | Yes | Not needed | High (same optimization) | High (step-by-step clarity) |

**Observations**

* All **three tools generated correct, efficient, and executable code** on first attempt.
* **ChatGPT**: Concise, best for professionals.
* **Gemini**: Most **readable & modular** (helper function).
* **Claude**: More **explicit and beginner-friendly** style.

**Conclusion:**

* For **production-ready coding** → **ChatGPT** is most efficient.
* For **learning & clarity** → **Claude** is easiest to follow.
* For **balanced readability & modular design** → **Gemini** is best.

**Task 4 – Usability & Interface Review**

**Objective:** Evaluate **user experience** beyond outputs.

**Comparison Table**

|  |  |  |  |
| --- | --- | --- | --- |
| Criteria | ChatGPT (GPT-5) | Google Gemini | Claude (Anthropic) |
| **Ease of Use** | Very easy, intuitive; widely adopted | Integrated with Google ecosystem, smooth but sometimes cluttered | Clean, distraction-free, simple UI |
| **Interface Design** | Professional, modern UI | Polished, integrated with Gmail/Docs/Drive | Minimalist, focuses on text |
| **Additional Features** | Plugins, file uploads, code interpreter, multimodal (text, images) | Strong multimodal (text, image, video), Google search integration | Document analysis (long context), supports large inputs |
| **Pricing / Free Limits** | Free tier + Plus subscription; good balance | Free with Google account; Pro plan available | Free limited use, Pro for extended context |

**150–200 Word Review**

After testing all three tools, it is clear that **ChatGPT, Gemini, and Claude each shine in different areas of usability**.

**ChatGPT (GPT-5)** offers the most **versatile ecosystem**. Its ease of use, professional design, and additional features like plugins, file uploads, and multimodal support make it a powerful assistant. The balance between free and paid tiers is fair, and it is particularly strong for **general writing and coding**.

**Google Gemini** stands out for **multimodal capabilities** (text, images, video) and seamless **integration with Google services** such as Gmail, Docs, and Drive. This makes it highly suitable for **research and contextual tasks**, though the interface may feel slightly cluttered compared to others.

**Claude** provides a **minimalist, distraction-free interface** with excellent long-context handling, making it ideal for reading and analyzing lengthy documents. However, it lacks the advanced ecosystem features of ChatGPT and Gemini.

**Recommendation:**

* For **general writing** → ChatGPT
* For **coding** → ChatGPT
* For **research and long document analysis** → Gemini and Claude