
UNVEILING AI TOOLS

Manjunatha KH*¹

*¹Department Of MCA, Brindavan College Of Engineering, Bangalore, Karnataka, India.

ABSTRACT

Artificial Intelligence (AI) tools, such as ChatGPT, Google BARD, and Genie AI Chatbot, have revolutionized industries through their advanced natural language processing and machine learning capabilities. This paper presents a comprehensive comparison of these AI tools, focusing on parameters such as response time, flexibility, robustness, accuracy of results, and prompt engineering. Through systematic evaluation and analysis, this study provides valuable insights to assist organizations and developers in selecting the most suitable tool for their specific needs. The findings aim to optimize AI implementation, promote healthy competition, and drive advancements in the field, leading to more efficient and effective AI solutions.

Keywords: AI Tools, Natural Language Processing, Large Language Models, Machine Learning, Response Time, Accuracy.

I. INTRODUCTION

Artificial Intelligence (AI) tools have transformed industries by addressing significant challenges that organizations faced before their development. Manual processes struggled to handle complex queries, resulting in slow response times and hindering prompt interactions. Moreover, limited flexibility and robustness made it difficult to adapt to diverse scenarios effectively. This research paper presents a comprehensive comparison of three prominent AI tools: ChatGPT, Google BARD, and Genie AI Chatbot. The evaluation methodology focuses on essential parameters such as response time, flexibility, robustness, accuracy of results, and prompt engineering. By analyzing the response time of each tool, the efficiency in generating prompt answers is measured. The flexibility of the tools is assessed through the exposure to various query types, including complex and context-dependent questions. Stress testing the tools under different scenarios evaluates their robustness in handling diverse situations. The accuracy of results is determined by comparing the relevance and correctness of the answers provided by each tool.

The comparison study aims to provide valuable insights for organizations and developers, allowing decision-makers to select the most suitable tool based on their specific requirements and constraints. Understanding the strengths and weaknesses of each tool in terms of response time, flexibility, robustness, accuracy of results, and prompt engineering empowers organizations to optimize their AI implementation. Additionally, this research promotes healthy competition in the field, fostering advancements that lead to more efficient and effective AI solutions.

II. METHODOLOGY

The methodology for this research paper follows a systematic approach to compare AI tools in natural language processing. It involves collecting relevant literature, designing experiments, stress-testing the tools, evaluating result accuracy, and analyzing prompt engineering strategies. This structured methodology ensures an objective and comprehensive analysis, providing valuable insights for organizations and developers.

A. Selection of AI Tools

The first step in the methodology is to select the AI tools for comparison. In this case, the chosen AI tools are ChatGPT, Google BARD, and Genie AI Chatbot. These tools have been selected based on their prominence in the field of natural language processing and their widespread adoption in various industries.

B. Definition of Evaluation Parameters

The next step involves defining the evaluation parameters to assess the performance of the AI tools. The parameters include response time, flexibility, robustness, accuracy of results, and prompt engineering. Response time refers to the duration taken by each tool to generate responses.

Flexibility assesses the ability of the tools to handle various query types, including complex and context-dependent questions. Robustness involves stress-testing the tools under different scenarios. Accuracy of results

is determined by comparing the relevance and correctness of the answers provided. Prompt engineering evaluates the effectiveness of the tools in generating prompt and contextually appropriate responses.

C. Data Collection

To conduct a comprehensive evaluation, a diverse dataset of queries and scenarios should be collected. This dataset should cover a wide range of query types, complexities, and context-dependent questions. The dataset can be collected from various sources, including real-world customer interactions, industry-specific use cases, and simulated scenarios.

D. Performance Evaluation

The AI tools are subjected to the collected dataset to evaluate their performance based on the defined parameters. The response time for each tool is measured and compared. The flexibility of the tools is assessed by analyzing their ability to handle different query types effectively. The tools are stress-tested under various scenarios to evaluate their robustness. The accuracy of results is determined by comparing the relevance and correctness of the answers provided by each tool. Prompt engineering is evaluated by examining the quality and contextuality of the responses generated.

E. Data Analysis

The collected data from the performance evaluation is analyzed to draw meaningful insights. Statistical analysis, such as calculating average response times and accuracy rates, can be performed. The strengths and weaknesses of each tool are identified based on the evaluation parameters. The analysis should focus on identifying patterns, trends, and significant differences among the tools.

F. Result Interpretation

The results of the evaluation and data analysis are interpreted to provide meaningful conclusions. The performance of each tool is assessed based on the evaluation parameters. The strengths and weaknesses of the tools are discussed, highlighting their suitability for different use cases and requirements. The interpretations should be supported by data and examples from the evaluation process.

G. Discussion and Conclusion

The findings from the evaluation and interpretation are discussed in the context of the research objectives. The implications of the results for organizations and developers are discussed, including the selection of the most suitable AI tool based on specific requirements and constraints. The limitations of the evaluation methodology are acknowledged, and future research directions are suggested. The research paper concludes by summarizing the key findings and their significance in advancing AI implementations.

III. MODELING AND ANALYSIS

Depending on the intricacy of the inquiry, ChatGPT usually produces results in a matter of seconds. Google BARD is well-known for its rapid reaction times; it is frequently compared to ChatGPT.

Adaptivity:

ChatGPT: Extremely adaptable, having the ability to grasp context and handle a broad range of topics and sophisticated queries.

Google BARD: Exceptionally adaptable and capable of managing a wide range of context-dependent queries. Though it occasionally lags behind the other two in processing exceedingly complex or highly technical queries, the Genie AI Chat bot offers a great deal of flexibility & solidity.

ChatGPT: Exhibits reliable operation under a range of conditions, such as stress testing with a large number of inquiries.

Google BARD: Similarly resilient, keeping accuracy and performance even under pressure.

The performance of the Genie AI chatbot is generally stable, however it may exhibit fluctuations in situations of high stress.

IV. RESULTS AND DISCUSSION

To assess the performance of the AI tools (ChatGPT, Google BARD, and Genie AI Chatbot), a comprehensive set of experiments was conducted. For each parameter, specific inputs were designed to evaluate the capabilities of the tools. For response time evaluation, 10 different queries of varying complexity and length were provided to

each tool to measure the time taken for response generation. In terms of flexibility, a diverse range of query types, including complex and context-dependent questions, were used to test the tools' adaptability and ability to understand different query formats.

To evaluate robustness, challenging inputs such as ambiguous questions and language variations were used to assess the tools' performance in generating accurate and meaningful responses

Accuracy of results was assessed by comparing the generated answers to expected or ground truth answers for 10 carefully selected inputs covering a wide range of topics. Prompt engineering was evaluated by analyzing the coherence and contextually appropriate nature of the responses generated by each tool.

The Results obtained for the experiments are tabulated below

Table 1: Comparison table

Parameters	ChatGPT	BARD	Genie AI
Avg Response Time(sec)	1.5	10	6
Robustness	82%	74%	70%
Flexibility	High	Moderate	Moderate
Accuracy	90%	75%	60%
Prompt Engineering	Excellent	Fair	Fair

V. CONCLUSION

In Conclusion, ChatGPT, with its sophisticated language model, excels in various parameters such as response time, flexibility, robustness, accuracy of results, and prompt engineering. It demonstrates a high level of flexibility in handling diverse query types and exhibits good prompt engineering. However, compared to the other tools, ChatGPT may have some limitations. For instance, it lacks direct internet access and the ability to provide real-time references, which may be essential for certain applications.

On the other hand, Google BARD offers the unique advantage of direct internet access, allowing it to provide references and information from real-time sources. Furthermore, BARD has recently introduced a feature to upload images, expanding its functionality. Although still under development as of February 2023, this image upload feature shows promising potential for enhancing user experiences and enabling more comprehensive interactions.

Genie AI stands out with its advanced features, including the ability to upload images, URLs, and PDF documents. It goes a step further by providing accurate summarizations of the uploaded content. Its capability to summarize information with high accuracy adds value for users seeking quick and reliable insights. However, a limitation lies in its limit of free chat interactions, as users are prompted for subscription after a certain threshold.

While ChatGPT excels in various aspects, Google BARD and Genie AI offer unique features and functionalities that cater to specific needs. Organizations and developers must carefully assess their requirements to determine the most suitable tool for their intended applications.

VI. REFERENCES

- [1] Hill-Yardin, Elisa L., et al. "A Chat (GPT) about the future of scientific publishing." *Brain Behav Immun* 110 (2023): 152-154.
- [2] Biswas, Som S. "Potential use of chat gpt in global warming." *Annals of biomedical engineering* 51.6 (2023): 1126-1127.
- [3] Sallam, Malik. "ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns." *Healthcare*. Vol. 11. No. 6. MDPI, 2023.
- [4] McGee, Robert W. "What Will the United States Look Like in 2050? A ChatGPT Short Story." *A Chatgpt Short Story* (April 8, 2023) (2023).
- [5] Dao, Xuan-Quy. "Performance Comparison of Large Language Models on VNHSGE English Dataset: OpenAI ChatGPT, Microsoft Bing Chat, and Google Bard." *arXiv preprint arXiv:2307.02288* (2023).

-
- [6] Agarwal, Mayank, Priyanka Sharma, and Ayan Goswami. "Analysing the Applicability of ChatGPT, Bard, and Bing to Generate Reasoning-Based Multiple-Choice Questions in Medical Physiology." *Cureus* 15.6 (2023).
- [7] Rudolph, Jürgen, Shannon Tan, and Samson Tan. "War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education." *Journal of Applied Learning and Teaching* 6.1 (2023).
- [8] Rahaman, Md Saidur, et al. "The AI race is on! Google's Bard and OpenAI's ChatGPT head to head: an opinion article." Mizanur and Rahman, Md Nafizur, *The AI Race is on* (2023).
- [9] Ram, Bal, and Pratima Verma. "Artificial intelligence AI-based Chatbot study of ChatGPT, Google AI Bard and Baidu AI." *World Journal of Advanced Engineering Technology and Sciences* 8.01 (2023): 258261.
- [10] Chomsky, Noam, Ian Roberts, and Jeffrey Watumull. "Noam Chomsky: The False Promise of ChatGPT." *The New York Times* 8 (2023).
- [11] Caramancion, Kevin Matthe. "News Verifiers Showdown: A Comparative Performance Evaluation of ChatGPT 3.5, ChatGPT 4.0, Bing AI, and Bard in News Fact-Checking." *arXiv preprint arXiv:2306.17176* (2023).
- [12] West, Darrell M. "Comparing Google Bard with OpenAIs ChatGPT on political bias, facts, and morality." (2023).
- [13] Rudolph, Jürgen, Shannon Tan, and Samson Tan. "War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education." *Journal of Applied Learning and Teaching* 6.1 (2023).
- [14] Fraiwan, Mohammad, and Natheer Khasawneh. "A Review of ChatGPT Applications in Education, Marketing, Software Engineering, and Healthcare: Benefits, Drawbacks, and Research Directions." *arXiv preprint arXiv:2305.00237* (2023).
- [15] Crawford, Joseph, et al. "Artificial Intelligence and Authorship Editor Policy: ChatGPT, Bard Bing AI, and beyond." *Journal of University Teaching & Learning Practice* 20.5 (2023): 1.
- [16] Hoppner, Thomas, and Luke Streatfeild. "Chatgpt, bard & co.: An introduction to ai for competition and regulatory lawyers." *An Introduction to AI for Competition and Regulatory Lawyers* (February 23, 2023) 9 (2023).
- [17] Kondapaneni, Rishil. "Using AI to Predict NBA Player Statistics ChatGPT vs Google Bard." (2023).
- [18] Patnaik, Sourav S., and Ulrike Hoffmann. "Comparison of ChatGPT vs. Bard to Anesthesia-related Queries." *medRxiv* (2023): 2023-06.