International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

Understanding AI Tool Preferences and Usage among Science Students in the Chalisgaon Region

Milind Bildikar

B. P. Arts, S. M. A. Science and K.K.C. Commerce College Chalisgaon, M.S., India

Abstract: ***Artificial intelligence (AI) has rapidly reshaped the academic behaviors of science students. This research investigates the application of AI technology by science students in Chalisgaon. Usage data were examined for ten notable AI products, specifically ChatGPT, Google Gemini, Perplexity, Remini, Grok, Dreamina, Character AI, Talkie AI, Deep Seek, and Quillbot. The average usage levels demonstrated considerable variation, indicating that students exhibited heterogeneous preferences and rates of adoption. Artificial intelligence (AI) has rapidly reshaped the academic behaviors of science students. This research investigates the application of AI technology by science students in Chalisgaon. We utilized one-way ANOVA and Tukey's HSD for our analysis. Usage data were examined for eleven notable AI products, specifically ChatGPT, Google Gemini, Perplexity, Remini, Grok, Dreamina, Character AI, Talkie AI, DeepSeek, and Quillbot. The average usage levels demonstrated considerable variation, indicating that students exhibited heterogeneous preferences and rates of adoption. This study provides important insights into the current trends of AI adoption among science students and emphasizes the need for proper awareness and guidance to ensure the effective and ethical use of AI tools in education.***

Keywords: Artificial intelligence, AI tools, One-way ANOVA, Tukey's Honestly Significant Difference etc.

1. Introduction

Artificial intelligence (AI) is a new technology that is changing the way students get, process, and use information in school. AI-based technologies have become more popular with students in the last few years, especially in scientific classes where analytical thinking, problem-solving, and putting together knowledge are important. These tools have a lot of different uses, such as generating text, analysing data, summarising content, and helping with interactive learning. This makes academic work more efficient and productive (Luckin et al., 2016; Holmes et al., 2019). ChatGPT and Google Gemini are two AI tools that have gotten a lot of attention because they are easy to use and feature sophisticated natural language processing capabilities. Tools like Perplexity AI, QuillBot, and Character AI are also being used more and more for specialised academic activities, such finding information, rephrasing, and getting people involved in the process. The increasing availability of these technologies has transformed conventional educational methodologies, fostering self-directed and technology-enhanced learning (Zawacki-Richter et al., 2019). Even though many students use AI tools, there is a lot of variation in how they use them and how well they work. The choice of tools is influenced by factors such how easy they are to use, how accurate they are, how well they work, and how beneficial they seem (Dwivedi et al., 2023). Also, certain tools are made for broad use, while others are more specific, which makes them more or less popular in different academic settings.

In India, especially in semi-urban areas like Chalisgaon, students are still figuring out how to use AI tools. Few studies have investigated the use of AI technologies by science students in these countries and whether notable disparities exist in their usage patterns. To encourage the safe and responsible use of AI in education, it is important to understand these tendencies. Consequently, the current study seeks to assess and contrast the utilisation patterns of diverse AI technologies among science students in the Chalisgaon region. The study aims to discern the most favoured tools by examining variations in usage levels and to offer insights into the changing role of AI in academic learning.

2. Material and Method

A survey was conducted at the Arts, Science, and Commerce Graduate College of Chalisgaon in Jalgaon District from August 2025 to December 2025. For this study, 100 individuals from each class within the scientific faculty were selected. Fifty male students and fifty female students participated from each class selected for the investigation. The sample size for this study was calculated using the Yamane Taro method (1967), based on data acquired from questionnaires completed by BSc and MSc students. The survey comprised 10 questions. All enquiries concerning the use of AI applications encompass ChatGPT, Google Gemini, Perplexity, Remini, Grok, Dreamina, Character AI, Talkie AI, Deep Seek, and QuillBot. Compute the mean for each source based on the usage by F.Y. BSc, S.Y. BSc, T.Y. BSc, MSc-I, and MSc-II. Subsequently, arrange the data in decreasing order according to the arithmetic mean presented in the final column of the table. A mixed methods technique was utilised to investigate the study problem. Quantitative data was obtained using surveys. The survey participants were college students from Chalisgaon, aged 18 to 28 years. Random sampling was conducted.

3. Statistical analysis-

One-way ANOVA was performed in conjunction with Tukey's Honestly Significant Difference (HSD). Typically, a significance level (α) of 0.05 is used. The Q statistic measures the difference between the means of treatments.

Formulate Hypotheses

- Null Hypothesis: There is no significant difference in the mean usage of different AI tools among science students.

- Alternative Hypothesis: There is a significant difference in the mean usage of different AI tools among science students.

4. Observation

Table: Observation table shows percentage of source users

S.no	Sources	Percentage of user students of each class				
		F.Y.BSc. (%)	S.Y.BSc. (%)	T.Y.BSc. (%)	MSc -I (%)	MSc-II (%)
1	Google Gemini	40	50	80	90	90
2	ChatGPT	75	80	90	100	100
3	Perplexity,	26	30	50	70	70
4	Remini	14	30	35	35	37
5	Grok,	12	20	20	25	25
6	Dreamina	10	10	15	16	18
7	Character AI	30	30	50	55	78
8	Talkie AI	40	56	63	78	78
9	Deep Seek,	25	30	45	50	60
10	Quill bot.	05	08	10	15	80

Anova: Single Factor

Groups	*Count*	*Sum*	*Average*	*Variance*
Google Gemini	5	350	70	550
ChatGPT	5	445	89	130
Perplexity	5	246	49.2	443.2
Remini	5	151	30.2	88.7
Grok	5	102	20.4	28.3
Dreamina	5	69	13.8	13.2
Character AI	5	243	48.6	399.8
Talkie AI	5	315	63	257
DeepSeek	5	210	42	207.5
Quillbot	5	118	23.6	1007.3

ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	25900.98	9	2877.88667	9.20923733	2.17E-07	2.124029
Within Groups	12500	40	312.5			

5. Result and Discussion

The one-way analysis of variance (ANOVA) demonstrated a statistically significant difference among the ten assessed AI tools, showing substantial variety in their average performance or utilisation levels. The analysis yielded an F-value of 9.21 and a p-value of 2.17×10^{-7} ($p < 0.01$), indicating that at least one treatment mean significantly differed from the others. To investigate these differences further, various post hoc tests, including Tukey's Honestly Significant Difference (HSD), Scheffé, Bonferroni, and Holm tests, were utilised for pairwise comparisons. The Tukey HSD test revealed several significant differences, indicating that ChatGPT exhibited substantially higher mean values than Remini, Grok, Dreamina, and Quillbot ($p < 0.01$). Google Gemini demonstrated markedly superior performance compared to less effective tools like Grok and Dreamina ($p < 0.01$), while comparisons among mid-tier tools such as Perplexity, Character AI, DeepSeek, and Talkie AI were predominantly non-significant, suggesting similar performance within this category. The Scheffé test, being more conservative, identified fewer significant differences but affirmed that ChatGPT was significantly superior to several lower-performing tools, specifically Remini, Grok, Dreamina, and Quillbot ($p < 0.01$), and that Google Gemini exhibited significant differences from Grok and Dreamina ($p < 0.05$). Most additional comparisons under Scheffé were non-significant, indicating the presence of homogeneous groupings. The Bonferroni and Holm corrections corroborated these findings, indicating that ChatGPT exhibited substantial differences from various tools, including Remini, Grok, Dreamina, and Quillbot ($p < 0.01$), while Google Gemini also demonstrated significant disparities with these lower-performing tools ($p < 0.01$). Furthermore, certain comparisons, including ChatGPT versus Perplexity and Character AI, were statistically significant at $p < 0.05$. Comparative analyses revealed significant differences between Google Gemini and Remini, Grok, Dreamina, and Quillbot ($p < 0.01$), while no statistically significant differences were found between Google Gemini and ChatGPT, as well as other mid-tier tools, suggesting comparable performance within the higher and intermediate categories.

This study reveals statistically significant variances among AI tools, indicating diversity in their performance or usage trends. The substantial ANOVA result indicates that the tools are not equal, requiring post hoc analyses to determine particular group disparities.

In all post hoc analyses, ChatGPT consistently ranked as the highest-performing tool, demonstrating markedly superior mean values compared to several other tools. The persistent dominance observed in the Tukey, Scheffé, Bonferroni, and Holm tests enhances the credibility of this finding. The supremacy of ChatGPT can be ascribed to its sophisticated

natural language processing abilities, extensive training data, and wider application across various tasks, as indicated in recent research on big language models (OpenAI, 2023; Brown et al., 2020).

Google Gemini had robust performance, frequently securing the second position and displaying notable disparities relative to less effective technologies. Nonetheless, the disparities with ChatGPT were not statistically significant, indicating analogous performance levels across these two preeminent systems. This corresponds with current progress in multimodal and generative AI systems created by leading technology firms. The mid-tier group, comprising Perplexity, Character AI, DeepSeek, and Talkie AI, exhibited predominantly negligible variances among its members. This suggests a very uniform performance level throughout this group, maybe indicating analogous underlying systems or application domains. Comparable performance clustering has been noted in investigations of AI tools, wherein mid-range systems demonstrate overlapping capabilities (Zhao et al., 2023). Conversely, underperforming technologies like Remini, Grok, Dreamina, and Quillbot repeatedly exhibited markedly lower mean values, especially in comparison to ChatGPT and Google Gemini. These discrepancies may arise from enhanced specialised functionalities, restricted model capacity, or constrained application domains. Tukey HSD discovered the greatest number of significant differences among the statistical methods employed, but the Scheffé test recognised only the most substantial contrasts, indicative of its conservative approach. The Bonferroni and Holm techniques offer a compromise between sensitivity and Type I error control, with Holm demonstrating marginally more power than Bonferroni (Holm, 1979). The uniformity of principal findings across all methodologies strengthens the credibility of the conclusions.

The findings delineate a distinct hierarchy among the assessed AI tools, with ChatGPT and Google Gemini in the forefront, succeeded by a cohort of moderately performing tools, and a collection of lower-performing tools. These findings are significant for consumers and researchers in choosing suitable AI tools according to performance criteria. The one-way ANOVA revealed a statistically significant difference among the ten AI tools ($p < 0.01$). Post hoc analyses (Tukey, Scheffé, Bonferroni, and Holm) repeatedly demonstrated that ChatGPT exhibited considerably superior performance compared to many lower-performing tools.

References

[1] Artificial Intelligence related comparative studies (Zhao et al., 2023).
[2] Olaf Zawacki-Richter et al. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16(1), 39.
[3] OpenAI. (2023). *GPT-4 technical report*.
[4] Recent developments in generative AI systems (<PRIVATE_PERSON> et al., 2024).
[5] Rose Luckin et al. (2016). *Intelligence unleashed: An argument for AI in education*.
[6] Sture Holm. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.
[7] Tom B. Brown, et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
[8] Wayne Holmes, Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*.
[9] Yogesh K. Dwivedi et al. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on AI. *International Journal of Information Management*, 71, 102642.