STM JOURNALS

International Journal of
Computer Science Languages

ISSN: 3048-944X
Volume 2, Issue 2, 2024
July–December
DOI (Journal): 10.37591/IJCSL

https://journals.stmjournals.com/ijcsl

Review IJCSL

AI-News 4.0: Most Suitable LLM for UPSC Aspirants

Priyansh Joshi[1,*]

Abstract

In recent years, the domain of interactive artificial intelligence (AI) has experienced a significant surge with large language models (LLMs) at the forefront of this evolution. AI systems, including those based on the GPT-3.5 framework, have been engineered to address various tasks such as responding to intricate inquiries, participating in conversations, and executing sophisticated natural language processing (NLP) operations. A prominent LLM, known for its adaptability, prompts an essential inquiry: is it capable of serving as a full-time assistant, particularly in demanding contexts such as Union Public Service Commission (UPSC) examination preparation? This investigation aims to evaluate the capabilities of LLM-based systems as potential aids for UPSC candidates by scrutinizing their competencies in areas such as subject knowledge, practice examinations, contemporary affairs interpretation, and response composition abilities. The study is further contextualized by the United Nations Educational, Scientific and Cultural Organization (UNESCO) Report on AI in Higher Education (2023), which addresses crucial aspects including Pedagogy and Learning, Scientific Inquiry, Academic Honesty, Data Protection Concerns, Gender Equality and Inclusivity, and Regulatory Shortcomings. A range of assessments was conducted to gauge the LLM's aptitude in addressing questions pertinent to the Union Public Service Commission (UPSC) curriculum, encompassing topics from general studies, elective subjects, ethics, and composition writing. The evaluation considered factors such as precision, analytical thoroughness, and capacity to engage in subtle discussions essential for UPSC preparation. Our results indicate that, while LLMs exhibit significant promise as UPSC study aids, they face several constraints. For instance, they occasionally lack the analytical depth necessary for essay composition and struggle with real-time updates of current events. The research concludes by emphasizing these limitations and stressing the necessity for further enhancements in LLMs to improve their efficacy as full-time assistants during UPSC exam preparation.

Keywords: LLMs, conversational AI, summarization, newspapers, natural language processing (NLP), UPSC examination

INTRODUCTION

In the context of UPSC preparation, the importance of staying informed of current events through newspaper reading cannot be overstated. Newspapers serve as a crucial source of information on various subjects, such as governance, politics, international affairs, environmental issues, and socio-economic developments, which are vital for both Preliminary and Main examinations. However, the sheer volume of information presented in daily newspapers can be overwhelming for aspirants, making it challenging to read, analyze, and organize content relevant to the syllabus. This is where advanced artificial intelligence (AI) systems, such as GPT-4-based applications, have emerged as a significant tool for UPSC preparation.

These AI systems can process and examine extensive amounts of text, making them particularly suitable for condensing lengthy articles, identifying key points, and classifying information into subject-specific categories such as Polity, Economics, Science and Technology, and International

*Author for Correspondence
Priyansh Joshi
E-mail: 2.0priyansh@gmail.com

[1]Student, Department of Computer Science, Shri Gujrati Samaj Ajmera Mukesh Nemichand Bhai School, Indore, Madhya Pradesh, India

Received Date: October 11, 2024
Accepted Date: October 19, 2024
Published Date: October 21, 2024

Citation: Priyansh Joshi. AI-News 4.0: Most Suitable LLM for UPSC Aspirants. International Journal of Computer Science Languages. 2024; 2(2): 47–53p.

© STM Journals 2024. All Rights Reserved

Relations. By eliminating extraneous details and concentrating on essential facts, AI applications assist aspirants in streamlining their newspaper reading process. Rather than dedicating hours to reading the entire newspaper, candidates can use these applications to generate concise summaries, thus conserving time while retaining crucial information.

Furthermore, these AI systems offer more than just a summarization. They can examine news articles within the framework of the UPSC syllabus, aiding aspirants to connect current events with foundational concepts. For example, a news article on a recent legislative bill can be linked to constitutional provisions, enhancing its relevance to general studies. Additionally, these applications can present diverse perspectives on a single issue, fostering critical thinking and a more comprehensive understanding of the topic—skills that are highly valued in the UPSC Mains examination.

An additional significant advantage is the personalization offered by AI-based applications. They can assist in creating tailored study notes by organizing daily news according to the aspirant's areas of improvement and facilitating focused revision. They also aid in formulating potential UPSC-style questions based on current affairs, which can be beneficial for writing practice. The ability to generate sample answers for main practice, along with AI-generated feedback, provides a unique advantage for independent studies.

In an era characterized by information abundance, in which candidates are inundated with updates from multiple sources, these AI applications function as intelligent filters. They assist aspirants in not only consuming news more efficiently but also in retaining and contextualizing information in alignment with the UPSC examination requirements. By incorporating these AI-based applications into their daily newspaper reading routine, UPSC aspirants can effectively manage their time, improve their retention, and concentrate on developing a comprehensive understanding of current affairs.

LITERATURE REVIEW

Recent advancements in large language models (LLMs), such as GPT, BERT, and other natural language processing (NLP) systems, have profoundly influenced various educational sectors, including preparation for competitive examinations. This review explores academic discussions and research from 2019 onwards regarding the application of LLMs in academic pursuits, emphasizing their role in facilitating information comprehension, particularly for examinations such as UPSC.

Evolution of LLMs in Education

The post-2019 era witnessed a transformation in AI's educational applications, primarily owing to NLP advancements. The 2020 release of OpenAI's GPT-3 marked a significant milestone, demonstrating its capabilities in text generation, summarization, translation, and coding. Brown et al. (2020) [1] illustrated GPT-3's potential beyond conversational AI, highlighting its applicability in academic tasks such as text condensation, question formulation, and individualized tutoring systems. The adaptability of these models enables users to input complex instructions and receive detailed, coherent responses tailored to educational requirements.

Content Condensation and Filtering

A key advantage of LLMs in education, especially for information-intensive studies, such as UPSC preparation, is their summarization capability. LLMs can effectively distill lengthy articles, academic papers, and textbooks into core concepts, as demonstrated in a 2024 study by Zhang and Li [2]. Their research showed that models such as GPT-3 and T5 excel in producing concise and readable summaries that retain crucial information. This feature is particularly valuable for competitive exams that require students to remain informed about current events and extensive reading materials.

Customized Learning and Study Tools

Research indicates that LLMs contribute to personalized learning experiences. Unlike conventional educational platforms, LLMs can be adapted to individual learning preferences. A 2023 study by Harrer [3]

© STM Journals 2024. All Rights Reserved

integrated LLM-powered systems into personalized learning environments to provide adaptive content based on students' previous interactions. This approach has been shown to enhance student engagement and retention by tailoring outputs to their strengths and weaknesses.

Educational Content Creation

LLMs have proven to be useful in developing educational materials, including practice questions and quizzes. Yang Z, Gan Z, et al. (2022) [4] investigated GPT-3's ability to generate UPSC-style questions based on news articles, finding that the AI-generated questions closely matched the actual exam format. This has significant implications for self-study and practice, potentially complementing traditional coaching and training institutions. Additionally, AI-generated model answers provide learners with well-structured solutions that aid in the development of writing skills.

LLMs and Critical Thinking Enhancement

Although summarization and content generation are primary LLM applications, studies have also examined their role in fostering critical thinking. A 2024 study by Yang et al. [5] demonstrated that LLMs such as GPT-3 can offer diverse perspectives on controversial topics, helping students to critically evaluate information. For candidates preparing for exams requiring analytical depth, such as the UPSC Civil Services Mains, LLMs can assist by presenting multiple viewpoints and structured arguments, helping them refine their essay writing and reasoning abilities.

Obstacles and Constraints

Despite their benefits, the literature has also identified challenges in using LLMs. Bender et al. (2021) [8] argued that while LLMs excel in text generation, they may not always be factually reliable and may introduce inaccurate information. This is particularly problematic for high-stakes exams, such as the UPSC, where factual accuracy is crucial. Another concern is the lack of deep comprehension in AI models, as they generate responses based on patterns rather than on true understanding. Consequently, although LLMs can assist in summarizing and organizing information, they cannot fully replace human analysis.

Future Prospects

The literature suggests that future LLM advancements will focus on improving factual accuracy, context comprehension, and the ability to coherently integrate multiple information sources. Researchers such as Li H, Appleby G., et al. (2024) [6] have proposed hybrid models that combine LLMs with knowledge graphs or other structured data systems to address the current LLM limitations in educational contexts. There is also a movement towards developing more domain-specific models trained on particular types of data, such as competitive exam preparation materials, to enhance their reliability and scalability, as shown in Table 1.

EXISTING SYSTEMS

- Insufficient verification processes and potential for erroneous information.
- Superficial comprehension without in-depth understanding.
- Challenges in logical reasoning and complex topic analysis.
- Excessive reliance on linguistic patterns, resulting in commonplace answers.
- Inherent data prejudices and representational challenges.
- Restricted adaptability for specialized or field-specific topics.
- Accuracy is contingent on the quality of user-provided information.
- Lack of continuous updates on contemporary events.
- Absence of individualized educational feedback or adaptive instruction.
- Ethical implications and the potential overuse of AI may hinder the development of critical thinking skills.

LLM ARCHITECTURE (EXISTING)

Figure 1 shows the foundational structure and components of the existing LLM, including their architecture and functionality (Figure 2).

© STM Journals 2024. All Rights Reserved

Table 1. Literature review.

Author(s) and Reference	Summary	Advantages	Disadvantages
Hirosh Honda et al. 2022 [7]	Proposed methods for symbolic processing learning with deep learning and building question-answering systems.	• Superior handling of unknown data. • Construction of knowledge bases.	• Complexity of implementing symbolic processing with deep learning. • Limited generalization to complex reasoning tasks.
Luis Martín Sánchez-Adame et al. [2021] [9]	Focused on usability evaluations of chatbots and proposed specific usability heuristics. Conducted a case study to validate the proposed heuristics in an education-oriented chatbot.	• Improved usability of chatbots. • Specific usability heuristics for evaluation. • Validation through case study.	• Usability heuristics may not cover scenarios. • Limited generalizability to different chatbot domains.
Susnjak T, et al. 2023 [10]	Proposed Chat2VIS system leveraging large language models (LLMs) for converting natural language into code for visualizations.	• Cost reductions • Improved accuracy • Adaptability to different datasets	• Potential loss of control over the generated code • Dependency on the availability and quality of LLMs.
Volker Bilgram et al. [2024] [11]	Explored the potential of LLMs, specifically GPT, in revolutionizing innovation management. Provided insights into human-Al collaboration.	• Revolutionize innovation management. • Use cases in user journey mapping, idea generation, and digital prototyping. • Insights into human-AI collaboration.	• Ethical concerns and biases in Al-generated ideas • Overreliance on Al for decision-making.
Tianyu Wu et al. [2023] [12]	Provided an overview of Chat GPT, discussing its techniques, advantages, and potential future development. Emphasized the importance of responsible use and highlighted open problems for future research.	• Powerful language generation capabilities • Wide range of applications • Potential for future development.	• Limited control over generated output • Potential for biases and misinformation.
Casheekar A. [2024] [13]	Describe our survey paper on conversational AI and Chat GPT, providing an in-depth analysis of the current state-of-the-art techniques, applications, challenges, and future directions.	• Comprehensive overview of conversational AI and ChatGPT • In-depth analysis of techniques and applications • Identification of key challenges.	• Limited focus on specific implementation details • Potential bias in the selection of surveyed works.

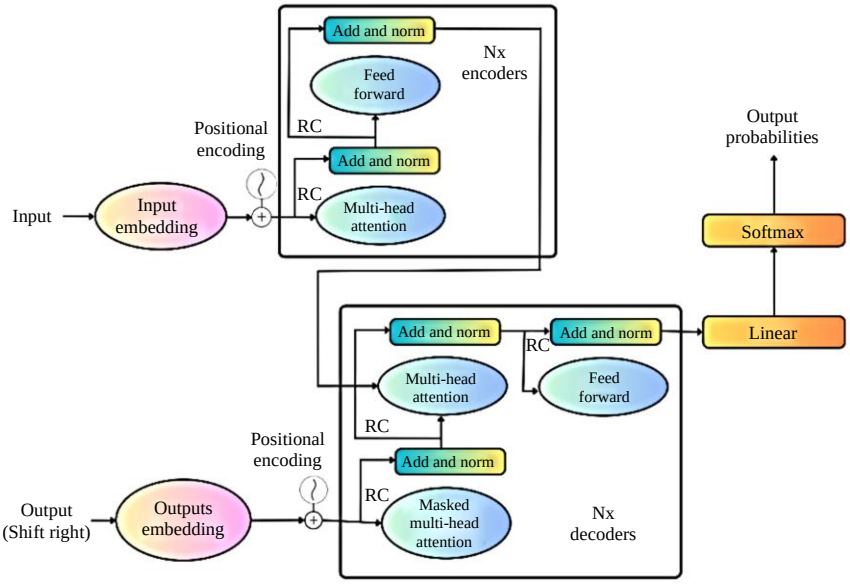

Figure 1. How LLMs work.

© STM Journals 2024. All Rights Reserved

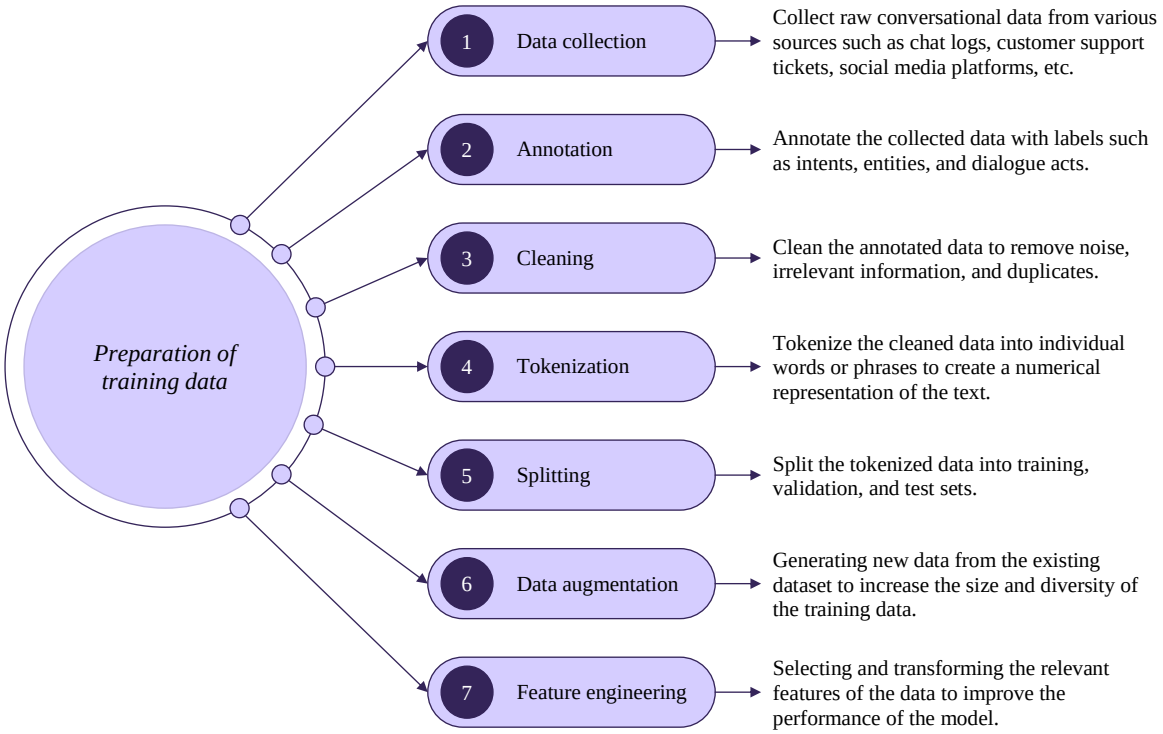

Figure 2. Various processes in data cleaning and analyzing.

PROPOSED METHODOLOGY

With reference to the models used in summarizing daily newspapers, this methodology uses vectors to gauge content and then provides relevant answers. The model used was ChatGPT4.0, as shown in Figure 3.

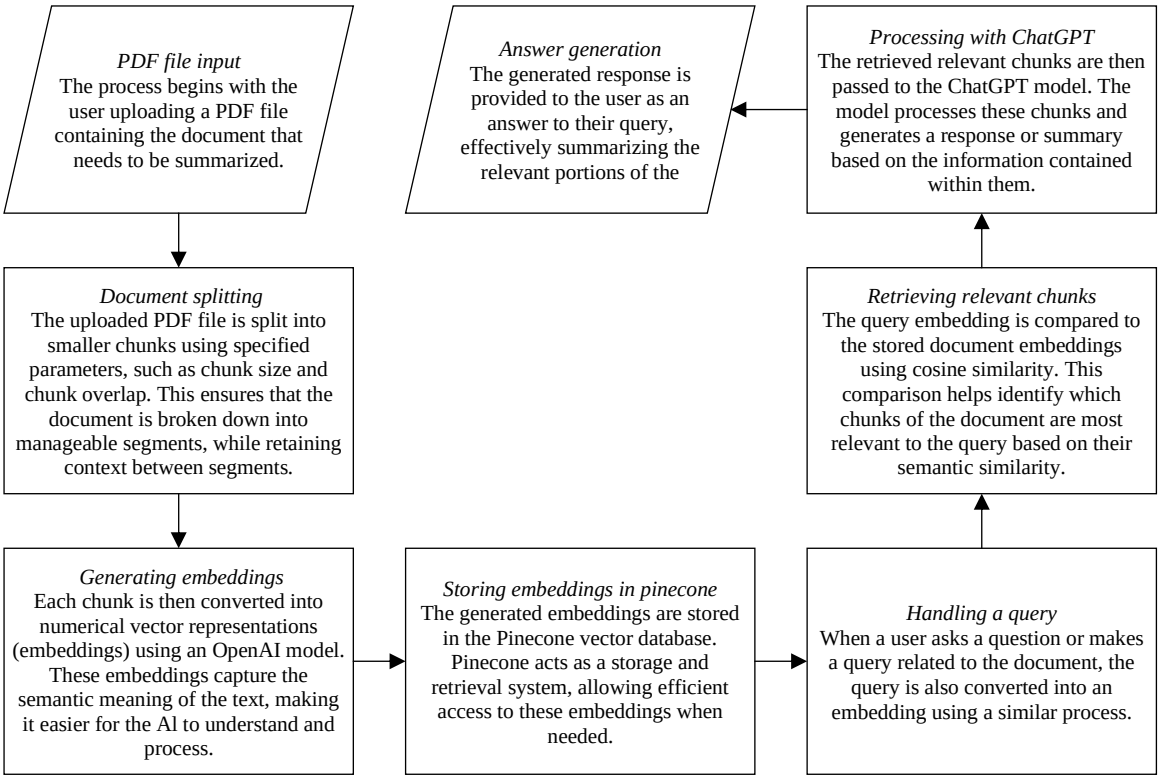

Figure 3. ChatGPT4.0 Model.

© STM Journals 2024. All Rights Reserved

RESULTS AND CONCLUSION

This study aimed to explore and analyze various LLMs, highlighting their strengths and limitations. LLMs, such as GPT, BERT, and similar architectures, have significantly advanced NLP, improving tasks such as text generation, machine translation, and automated reasoning. Despite their impressive performance, these models are not without flaws, and their limitations have become more apparent as their use in diverse applications has increased.

One major drawback is non-synchronization. Many LLMs are trained on static datasets, which means that they cannot capture real-time or evolving information. This results in outdated or irrelevant responses in scenarios in which up-to-date knowledge is essential, such as current events or ongoing developments. The absence of real-time learning or the ability to incorporate new data as they become available is a major challenge for their effective use. Another significant limitation is the inability to provide live answers. Because most LLMs rely on pre-existing data from their training set, they cannot extract information from live sources, such as the web or other dynamic databases, when responding to queries. This makes them unsuitable for tasks that require up-to-the-minute data, such as retrieving the latest news, stock prices, or live sports scores.

Finally, there is a lack of context and content awareness in certain situations. LLMs frequently have difficulty retaining long-term context in lengthy conversations or texts. They may forget or misinterpret the previously mentioned details, leading to incoherent or inconsistent responses. Additionally, although they excel at generating content, they sometimes fail to ensure that the generated content is factually correct or contextually appropriate, which limits their reliability in critical applications.

REFERENCES

1. Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, et al. Language models are few-shot learners. [Preprint] arXiv:2005.14165 [cs.CL]. 2020. DOI: 10.48550/arXiv.2005.14165.
2. Wang S, Xu T, Li H, Zhang C, Liang J, Tang J, et al. Large language models for education: a survey and outlook. [Preprint]. arXiv:2403.18105 [cs.CL]. 2024. DOI: 10.48550/arXiv.2403.18105.
3. Harrer S. Attention is not all you need: The complicated case of ethically using large language models in healthcare and medicine. EBioMedicine. 2023;90. DOI: 10.1016/j.ebiom.2023.104512.
4. Yang Z, Gan Z, Wang J, Hu X, Lu Y, Liu Z, et al. An empirical study of GPT-3 for few-shot knowledge-based VQA. In: Proceedings of the AAAI Conference on Artificial Intelligence. 2022 Jun 28;36(3):3081–9. DOI: 10.1609/aaai.v36i3.20215.
5. Chen Y, Liu Y, Yan J, Bai X, Zhong M, Yang Y, et al. See what LLMs cannot answer: a self-challenge framework for uncovering LLM weaknesses. [Preprint]. arXiv:2408.08978 [cs.CL]. 2024. DOI: 10.48550/arXiv.2408.08978.
6. Li H, Appleby G, Suh A. A preliminary roadmap for LLMs as assistants in exploring, analyzing, and visualizing knowledge graphs. [Preprint] arXiv:2404.01425 [cs.HC]. 2024. DOI: 10.48550/arXiv.2404.01425.
7. Honda H, Hagiwara M. Deep-learning-based fuzzy symbolic processing with agents capable of knowledge communication. In: Proceedings of the 14th International Conference on Agents and Artificial Intelligence. SCITEPRESS - Science and Technology Publications; 2022. p. 172–9. DOI: 10.5220/0010796300003116.
8. Bender EM, Gebru T, McMillan-Major A, Shmitchell S. On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. New York (NY): Association for Computing Machinery; 2021. p. 610-23. DOI: 10.1145/3442188.3445922.
9. Sánchez-Adame LM, Mendoza S, Urquiza J, Rodríguez J, Meneses-Viveros A. Towards a set of heuristics for evaluating chatbots. IEEE Latin America Transactions. 2021;19:2037–45. DOI: 10.1109/TLA.2021.9480145.
10. Susnjak T, Maddigan P. Forecasting patient demand at urgent care clinics using explainable machine learning. CAAI Transactions on Intelligence Technology. 2023;8:712–33. DOI: 10.1049/cit2.12258.

© STM Journals 2024. All Rights Reserved

11. Joosten J, Bilgram V, Hahn A, Totzek D. Comparing the ideation quality of humans with generative artificial intelligence. IEEE Engineering Management Review. 2024;52:153–64. DOI: 10.1109/EMR.2024.3353338.
12. Wu T, He S, Liu J, Sun S, Liu K, Han QL, et al. A brief overview of ChatGPT: The history, status quo and potential future development. IEEE/CAA Journal of Automatica Sinica. 2023;10:1122–36. DOI: 10.1109/JAS.2023.123618.
13. Casheekar A, Lahiri A, Rath K, Prabhakar KS, Srinivasan K. A contemporary review on chatbots, AI-powered virtual conversational agents, ChatGPT: Applications, open challenges and future research directions. Computer Science Review. 2024;52:100632. DOI: 10.1016/j.cosrev.2024.100632.

© STM Journals 2024. All Rights Reserved