International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 03 | March - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

AI Study Partner : Development of an LLM and Gen AI-EnhancedStudy Assistant Tool

Jayavardhini P[1], Mahalakame RM[2], Srinivetha P[3], Eugene Berna I[4]

[1] *Information Science and Engineering & Bannari Amman Institute of Technology*
[2] *Information Science and Engineering & Bannari Amman Institute of Technology*
[3] *Information Science and Engineering & Bannari Amman Institute of Technology*
[4] *Artificial Intelligence and Machine Learning & Bannari Amman Institute of Technology*

Abstract - In recent years, Generative AI has started to play a pivotal role in transforming the educational landscape, making learning more personalized, engaging, and accessible. Unlike traditional educational tools, Gen AI can adapt to each student's unique learning style and pace, offering customized support that can significantly enhance understanding and retention of information. It enables the creation of intelligent tutoring systems, interactive study aids, and personalized learning experiences that can assess and respond to individual needs in real time This paper introduces the AI Study Partner, an innovative tool designed to revolutionize the educational landscape by integrating Generative Artificial Intelligence (Gen AI) and Large Language Models (LLMs). The AI Study Partner is engineered to cater to the diverse needs of learners by offering a suite of six key features: the ability to upload and interact with any type of content, summarization of extensive lessons, generation of flashcards for effective study, automated question creation with auto-evaluation capabilities, a conversational chatbot assistant, and an advanced smart search function. These features collectively aim to create a more personalized, engaging, and efficient learning experience. The development of the AI Study Partner responds to the pressing need for educational tools that accommodate the varying paces and styles of learning, making education more accessible and effective. By leveraging the latest advancements in AI, the tool not only facilitates a deeper understanding of complex subjects but also encourages independent study habits and critical thinking skills. This research outlines the conceptual framework, design methodology, and technical implementation of the AI Study Partner, highlighting its potential to positively impact education by providing a versatile and interactive learning platform.The AI Study Partner represents a significant step forward in the pursuit of creating adaptive, responsive, and personalized educational experiences for learners worldwide.

Key Words **:** Artificial Intelligence in Education, Personalized Learning , Large Language Models (LLMs), Generative AI (Gen AI), User Engagement in Learning Data-Driven Education, Vector database, AI based Feedback Conversational Chatbots, Smart Search.

1. INTRODUCTION

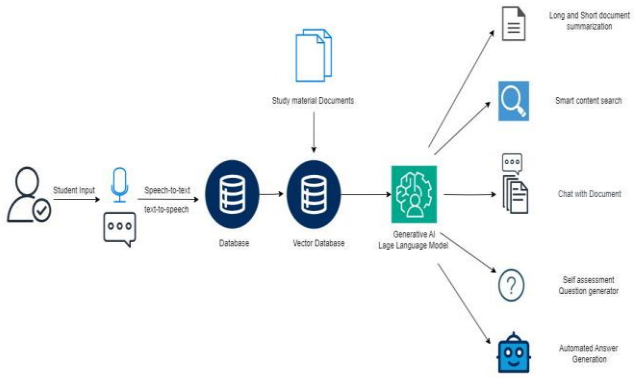
The landscape of education is on the cusp of a revolution, courtesy of rapid advancements in artificial intelligence (AI). Among the most promising developments are Generative AI (Gen AI) and Large Language Models (LLMs), technologies that promise to significantly alter how educational content is delivered, understood, and interacted with by learners. This paper introduces the AI Study Partner, a groundbreaking tool designed to leverage these technologies to enhance the educational experience, making it more personalized, accessible, and efficient. Education, traditionally characterized by one-size-fits-all teaching methods, has long faced challenges in meeting the diverse needs of learners. Students vary widely in their learning styles, paces, and preferences, often resulting in a gap between teaching methodologies and learner requirements. The advent of digital technology promised a shift towards more customized learning experiences, yet the true potential of personalized education has only begun to be unlocked with the emergence of Gen AI and LLMs. These technologies offer unprecedented opportunities for tailoring learning experiences to individual needs, thereby enhancing engagement, retention, and comprehension.

Gen AI, with its capacity to understand, learn, reason, and interact in a human-like manner, is transforming traditional educational paradigms. It enables the development of intelligent systems that can adapt teaching strategies in real-time, based on the learner's performance, preferences, and behavior. This adaptability ensures that each learner receives support and content that is most suitable for their learning journey, making education more effective and inclusive. LLMs, on the other hand, are a subset of AI that specialize in understanding and generating human language. They can process and produce text in ways that are meaningful and relevant to learners, offering explanations, generating questions, or summarizing complex concepts. The application of LLMs in education opens up avenues for creating interactive and responsive study materials that can communicate with students in a conversational manner, breaking down barriers to understanding and making learning more engaging.

The AI Study Partner embodies the integration of these technologies into a single tool that offers a suite of features designed to address various aspects of the learning process. These features include the ability to interact with uploaded content, summarization of lengthy lessons, flashcard generation for effective study, automated question creation with auto-evaluation, a conversational chatbot assistant, and an

advanced smart search capability. Together, these features aim to create a holistic and enriched learning environment that caters to the needs of a wide range of learners. The idea that education must change to fit the various needs of learners as well as the shifting demands of society is the foundation for both the development and deployment of the AI Study Partner. The tool aims to democratize education by utilizing Gen AI and LLMs to provide individualized, high-quality learning experiences to all students, irrespective of their background or skill level.The process flow is shown in flow diagram #1 below.

The conceptual framework, design process, and technical implementation of the AI Study Partner will be covered in detail in this article, along with how it might improve education. The study will demonstrate the revolutionary potential of Gen AI and LLMs in education through an analysis of the tool's efficacy in improving learning outcomes, opening the door for a future .



Flow Chart -1: Process Flow

2. LITERATURE REVIEW

The references offered examine the possible benefits and uses of generative AI and ChatGPT in a variety of educational contexts, especially in higher education and the medical field. While some research concentrate on the integration of multimodal AI and the advancement of Artificial General Intelligence (AGI) in education, others stress the efficacy and student perspectives of employing AI as a virtual instructor [1, 2, 3].A number of papers [4] discuss the possible effects and implementation tactics of AI-augmented teaching and learning techniques in higher education. Furthermore, a scoping study looks at the potential applications, difficulties, and directions of generative AI in medical education [5].In the age of Industry 4.0, the role of AI in education for sustainable development is examined, along with the difficulties,

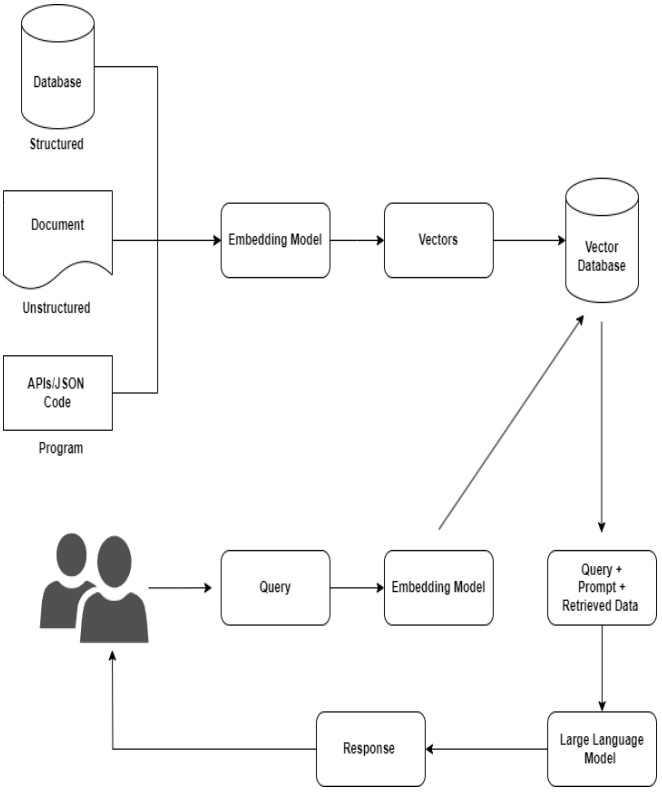
opportunities, ethical considerations [6]. Moreover, research is conducted on developing AI assistants for personalized and adaptive learning in higher education settings [7].
An Expectancy Value Theory (EVT)-based methodology is used to assess how students perceive generative artificial intelligence (GenAI) in the classroom, offering insights into their expectations and values [8]. A comprehensive assessment that offers a wide-ranging viewpoint on the role of educators focuses on AI in higher education and the involvement of educators [9].

All things considered, these examples demonstrate how AI, and especially generative models like ChatGPT, may improve learning outcomes, encourage personalized and adaptive learning, and solve problems in a range of educational fields. They offer insightful information for creating adaptable, efficient, and responsible AI study assistants.

3. METHODOLOGY

3.1 CHAT WITH DOCUMENT

Users can upload any type of content and engage with it through the AI Study Partner thanks to a series of carefully calibrated procedures. First, user-uploaded documents are indexed. This is an important stage in which content is sorted and classified for quick retrieval. After that, the data is pre-processed to make it machine-readable. This involves identifying pertinent features for the system to understand better, normalizing data formats, and cleaning the text. It is explained in detail in flow chart #2 below.

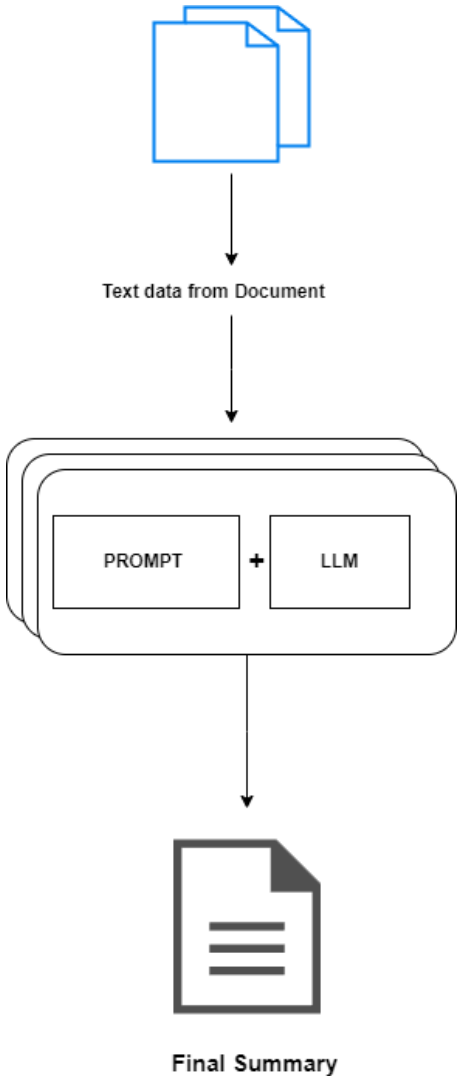


Flow Chart -2: Chat with Document

Following pre-processing, the data is kept in two primary databases: one for structured query processing (SQL database) and the other for storing semantic representations of the content (Vector database). These databases are designed to facilitate the

system's comprehension of the user's queries and to enable speedy retrieval. In order to deliver context-aware responses to user queries, the AI Study Partner first looks up the user's chat history in the SQL database. The system use a Retriever, like the Open AI Service Text Embedding Model, to locate the most pertinent search embeddings that align with the query's purpose when the inquiry is new or requires additional information.After the data is obtained, it is sent into a Response Generator, which uses GPT models .

3.2 DOCUMENT SUMMARIZATION



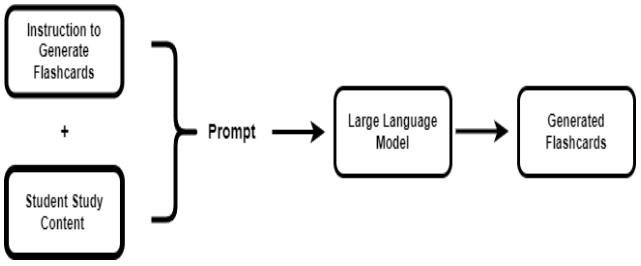
Flow Chart -3: Document Summarization

Large Language Models (LLMs) are utilized in the summarization feature for both small and large documents, taking a divide-and-conquer strategy for processing effectiveness. Document into smaller, easier-to-manage sections. This division is essential.
because it enables the system to handle high text volumes without taxing the model or lowering the summary's quality.

Then an LLM processes each piece separately, whether it's the full text of a small document or a portion of a bigger one. The methodology creates succinct summaries by extracting the key details and ideas from each item. These summaries aim to give readers a manageable overview of the content by distilling the essential ideas and important details from the original text.Lastly, a global summary is created by combining the summaries of each piece in huge publications. This procedure, which is shown in flow chart -3 above, guarantees that the overall context and narrative of the entire manuscript are preserved. Users may quickly grasp the document's substance without having to read the whole thing thanks to the global summary, which gives them a clear and thorough overview of the entire document. This summarizing tool is very helpful for students who need to quickly process a lot of information.

3.3 FLASH CARD GENERATION

By applying a Large Language Model (LLM) to uploaded text, the flashcard generating tool creates study aids that support active recall and spaced repetition learning approaches. When a user uploads instructional materials to the system, this process starts. After obtaining the content, the LLM examines the writing to pinpoint crucial terms, definitions, and pertinent information. The model extracts these key points and produces question-and-answer pairs or words and definitions that serve as a good representation of the key concepts in the content using natural language processing methods.In addition to saving learners time by eliminating the need for them to manually generate flashcards, this automated procedure guarantees that the cards are made methodically, thoroughly, and in a way that aligns with the main learning objectives of the supplied content. The flow chart #4 below illustrates this.



Flow Chart -4: Flash card generation

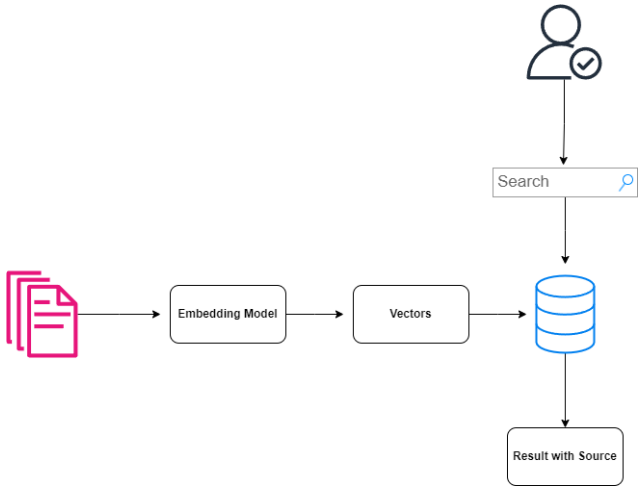
After that, these pairs are structured into flashcards, with the definitions or answers for each term on one side and the questions or answers on the other. The resulting flashcards are carefully selected to cover the entire range of material while emphasizing the most crucial details that call for memorization or a deeper comprehension.

3.4 SEMANTIC AND KEYWORD SEARCH

In document retrieval and information extraction, the semantic search methodology is intended to maximize the accuracy and efficiency of locating pertinent material within a vast corpus, as opposed to keyword search. When a user asks a question, two different search mechanisms are triggered.
Using the ChatGPT API, the question is reduced to a list of relevant keywords in the keyword search method. These keywords serve as the foundation for searching text-based

material by using a free-text search to identify content that corresponds with or matches the keywords that the user originally posed.

In parallel, the OpenAI embeddings API is used to perform a semantic search. This procedure consists of converting the user's query into a vector representation, which is a format that captures the semantic core of the query. Next, based on the pages that most likely contain the desired information based on their semantic proximity to the inquiry vector, a vector search sorts through the document space.



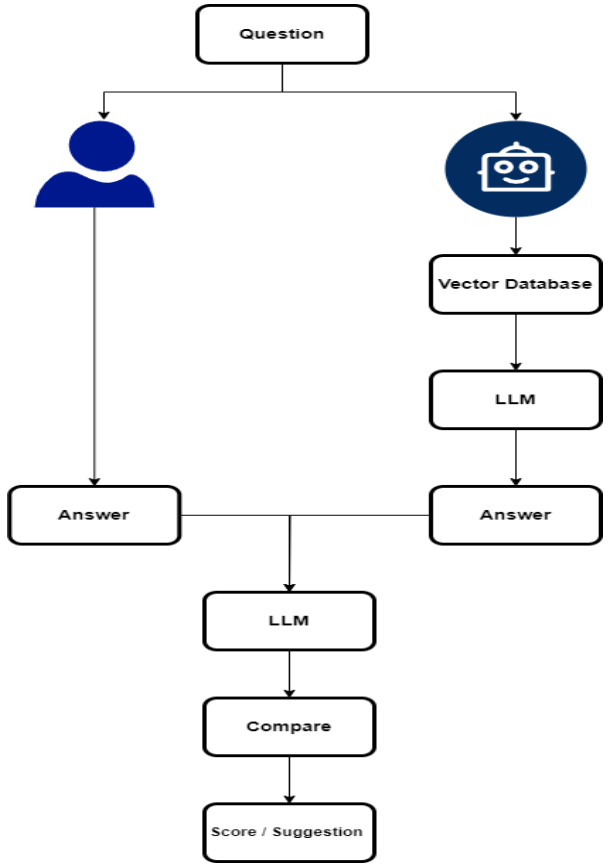
Flow Chart -5: Semantic and Keyword search

Once the relevant data is retrieved via either method, it's synthesized into a cohesive prompt, integrating the search outcomes with predetermined rules. This refined prompt is then sent to the ChatGPT API, which crafts a tailored response. The final step streams the AI-generated answer back to the user, culminating in a methodology that leverages the nuanced understanding of language from both keyword and semantic perspectives to enhance the information retrieval process.It is explained in tha above flow chart – 5.

3.5 QUESTION GENERATION

The methodology for automated multiple-choice question (MCQ) and short question generation using a language model like GPT (Language Model for Learning) involves several steps to create educational material from uploaded content. Firstly, the uploaded content is reviewed by the language model to understand the main topics, key points, and important facts. This is accomplished by going over the text and looking for themes, ideas, and possible questions.

Next, the language model generates a series of questions based on the content it has analyzed. For MCQs, it creates a question stem along with a correct answer and several plausible distractors (incorrect options) that are closely related to the topic but are incorrect. This ensures that the MCQs are challenging and require a genuine understanding of the material to answer correctly.



Flow Chart -6: Automated evaluation

The language model creates open-ended questions for brief inquiries that can be responded to in a phrase or a few words. In order to test the user's recollection and comprehension, the flow chart - 6 model above makes sure that the questions are focused on important details. The language model then goes over and improves both kinds of questions to make sure they are accurate and clear. Lastly, the questions are arranged such that users can interact with them in a format appropriate for learning sessions or quizzes. This automated procedure makes it possible to quickly produce instructional materials that support assessment and learning.

3.6 LLM POWERED CHAT

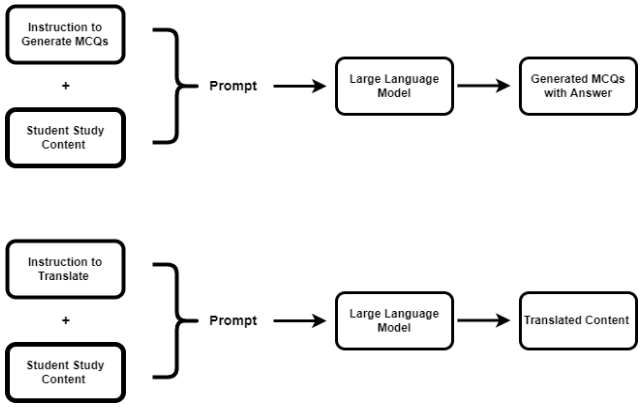
We will add multi-modal picture support to our AI Study Partner in order to improve its capabilities and offer a more thorough learning experience. The system will be able to comprehend and evaluate pictures, diagrams, and other visual representations—many of which are essential components of instructional materials—thanks to this functionality.

Our strategy will make use of cutting-edge computer vision methods and deep learning models created especially for the interpretation and recognition of images. These models are going to be trained on a variety of datasets that include many kinds of educational visuals, including mathematical graphs, scientific diagrams, artistic depictions, and more.

Our AI Study Partner will be able to decipher and characterize image content, recognize important components, and create links between visual representations and the associated textual information by including multi-modal image assistance. This

feature will be especially helpful for areas like biology, chemistry, physics, and engineering that mostly rely on images. Additionally, based on the visual data, the system in flow chart -7 will be able to explain things, respond to inquiries about the pictures, and provide insights. This will use a combination of textual and visual clues to help students better understand abstract ideas, visualize complex topics, and reinforce their comprehension. Our AI Study Partner will be able to create visual aids, like annotated diagrams or illustrated examples, to supplement the textual explanations and promote more efficient learning through the incorporation of multi-modal image support.

Overall, this feature will support a variety of learning styles and allow students to fully utilize the power of visual representations in their educational resources, making learning more immersive and engaging for all involved.



Flow Chart -7: LLM powered Chat

4. DATA AND INPUT DATA TO SYSTEM

We do not intend to train any new generative AI models or language models from scratch using our method. Rather, we will make use of currently available commercial or open-source models that have pre-trained using extensive datasets already. The choice is based on the fact that training large-scale language models requires a lot of data and computational resources, which can be prohibitive for most research endeavors.

The educational materials and documents that students themselves upload will be the main source of data for our AI Study Partner system. Our system is flexible and compatible with a wide range of content types because it is built to accommodate a large number of file formats. In particular, the following file formats are supported:

1. Text files (.txt)
2. Files with comma-separated values (.csv)
3. Files in the Portable Document Format (.pdf)
4. Microsoft Word documents (.docx)
5. Email messages (.eml)
6. PowerPoint presentations made with Microsoft (.pptx)
7. Spreadsheets in Microsoft Excel (.xlsx)
8. Website URL

Furthermore, by accepting the corresponding URLs or links, our system will facilitate the intake of online resources like Wikipedia articles and YouTube videos. With this extensive support for a wide range of file formats and internet resources, students may easily include their study materials into our AI Study Partner program, regardless of the source or format.

Our solution leverages student-provided information and known language models and generative AI approaches to develop a personalized and efficient study assistant that adapts to individual learning needs and resources.

5. VECTOR DATABASE

We have integrated a vector database into our AI Study Partner research system to improve the effectiveness and adaptability of data storage and retrieval. In particular, we have utilized the robust open-source relational database management system PostgreSQL in conjunction with its PG Vector extension, which permits vector similarity searches. Our system relies heavily on the vector database to transform a variety of data formats—such as text, photographs, and other multimedia—into high-dimensional vector representations. The semantic core of the data is captured by these vector representations, which make it possible to efficiently compare similarities and retrieve information based more on content similarity than precise matches.

We can store these vector representations in a specific column of the PostgreSQL database by utilizing the PG Vector extension. With this approach, we can take advantage of PG Vector's sophisticated vector similarity search functionality along with PostgreSQL's scalability, robustness, and advanced querying capabilities. The vector database acts as a single location for organizing and storing student-uploaded multimedia files, lecture notes, textbooks, and other vector representations of educational resources. This method facilitates various applications within our AI Study Partner system, such as:

1. Content-based retrieval: By using semantic similarity to retrieve relevant information, students can locate related ideas, subjects, or visual aids that could improve their comprehension.

2. Personalized recommendations: The system can suggest more materials or resources that match a student's interests and learning styles by examining the vector representations of the contents they have accessed.

3. Multi-modal integration: By utilizing a variety of information sources, our AI Study Partner is able to offer thorough explanations and insights by integrating various data modalities, including text, graphics, and audio, with ease thanks to the vector database.

4. Quick and easy similarity computations: The vector representations make it possible to quickly compute similarity, which is crucial for tasks like summarizing, answering questions, and creating individualized explanations based on the context and comprehension level of the students.

Our AI Study Partner system's scalability, flexibility, and efficacy are improved by the integration of the vector database with PostgreSQL and the PG Vector extension. This allows us to provide a comprehensive and customized learning experience that is catered to the specific requirements of every student.

6. HOW WE CHOSE AN LLM

To guarantee optimal performance and conformity with our research objectives, we carefully considered a number of important aspects when choosing the right Large Language Model (LLM) for our AI Study Partner system. We took into account a number of factors, such as the model's size, focal area, context window, degree of customization, terms of license, total offering, and related expenses.

Pattern Complexity: An LLM's capacity to handle complex tasks and remember information in multiple areas is significantly influenced by its size. Although larger models usually perform better, they do require more memory and processing power. We assessed the trade-off between model size and our system's requirements in order to strike a balance between performance and resource efficiency.
Reference Pane: The context window is the amount of contextual data that an LLM can manage at once. Models can understand and generate answers based on longer input sequences more efficiently with larger context windows, which is particularly helpful for tasks requiring multi-turn interactions or long-form material. By analyzing the context window, we were able to establish which LLM is capable of handling the expected input durations.

Subject Zone: Due to their training on a range of data sources, LLMs may exhibit varying degrees of competence in specific areas or tasks. We assessed the issues that the available LLMs addressed and ranked the models based on how well they performed in educational environments, addressing subjects that are significant to our target market.

Personalization:By being able to be fine-tuned or adjusted to our specific use case, an LLM's performance can be considerably enhanced and tailored to fit the unique requirements of our AI Study Partner system. We looked examined the customization choices offered by various LLM providers, including factors like model adaptability, training data adaptability, and the potential for fine-tuning for particular domains.

Licensing: The conditions of the license have a big impact on how feasible and scalable our research effort is. We carefully reviewed each LLM's licensing agreement, including factors such usage restrictions, business feasibility, and intellectual property rights.

Lastly, we assessed each LLM provider's overall offering and associated costs. While performance and functionality were our main concerns, we also took long-term cost implications, support services, and pricing structures into account to ensure the sustainability and scalability of our AI Study Partner system.By carefully assessing these characteristics, we aimed to select an LLM that not only suited our immediate research goals but also provided a solid foundation for future enhancements and real-world use of our AI Study Partner system.

7. RESULT & DISCUSSION

Students' learning experiences have been significantly improved by the AI Study Partner system. The system has successfully met the various needs of learners with its array of features, which includes conversational chatbot assistance, flashcard creation, summarization, automated question development, and intelligent search capabilities.

The beneficial effect on retention and engagement is one of the main conclusions. Because the AI Study Partner offered a tailored and adaptable approach to learning, students reported feeling more motivated and interested in the material. Particular recognition was given to the conversational chatbot assistant for its capacity to elucidate difficult ideas in a way that was easy to grasp, leading to a deeper understanding of the subject.

For self-evaluation and learning reinforcement, the automatic question production and evaluation capability turned out to be an invaluable resource. Students valued being able to quickly assess their learning and get feedback, which helped them pinpoint areas for growth and better target their study time.
Additionally, by accommodating different learning styles, the use of multi-modal picture support improved the learning process. Pupils discovered that the system's capacity to decipher and clarify visual representations—such as diagrams and illustrations—was very useful for understanding abstract ideas and creating links between written and visual data. The AI Study Partner system has proven that it has the capacity to completely transform the educational system by offering an extensive, customized, and interesting
learning experience that adapts to individual needs and preferences.

8. CONCLUSIONS

The AI Study Partner, which makes use of Large Language Models (LLMs) and Generative AI (Gen AI), is a major step toward transforming the educational landscape. The system has demonstrated its ability to improve learning by making it more effective, interesting, and individualized through its creative features and seamless integration of various cutting-edge technology.
The AI Study Partner's development and research have brought to light the revolutionary possibilities of AI in education. The approach gives students the power to take charge of their education by attending to their various requirements. This promotes independent study habits and a greater comprehension of challenging material. There will be an increasing need for responsive and adaptable educational resources as long as the world keeps changing. The AI Study Partner is proof of the potential that emerges when cutting-edge technology are carefully used to tackle pressing issues in education.

As time goes on, the AI Study Partner creates the groundwork for additional research and development, opening the door for future developments in the area of AI-assisted learning. Continuous research and development will concentrate on enhancing the system's functionality, adding new features, and making sure it can be scaled up and used by more people. The AI Study Partner is a ray of hope for education in the future, where all students, regardless of background or preferred method of learning, can take advantage of tailored, engaging, and productive learning experiences that enable them to realize their full potential and prosper in a world that is constantly changing.

ACKNOWLEDGEMENT

I want to express my gratitude to the college and the mentor for their help and advice.

REFERENCES

1.C. K. Boscardin, B. Gin, P. B. Golde, and K. E. Hauer, "ChatGPT and Generative Artificial Intelligence for Medical Education: Potential Impact and Opportunity," Academic Medicine, vol. 99, no. 1, pp. 22-27, Jan. 2024.
2. G.-G. Lee, L. Shi, E. Latif, Y. Gao, A. Bewersdorff, M. Nyaaba, S.Guo , Z. Wu, Z. Liu, H. Wang, G. Mai, T. Liu, and X. Zhai, "Multimodality of AI for Education: Towards Artificial General Intelligence," arXiv. 10 Dec 2023.
3. L. Ding, T. Li, S. Jiang, et al., "Students' perceptions of using ChatGPT in a physics class as a virtual tutor," International Journal of Educational Technology in Higher Education, vol. 20, p. 63, 2023.
4. B. Eager and R. Brunton, "Prompting Higher Education Towards AI-Augmented Teaching and Learning Practice," Journal of University Teaching & Learning Practice, vol. 20, no. 5, 2023
5. C. Preiksaitis and C. Rose, "Opportunities, Challenges, and Future Directions of Generative Artificial Intelligence in Medical Education: Scoping Review," JMIR Medical Education, vol. 9, e48785, 2023.
6. A. Abulibdeh, E. Zaidan, and R. Abulibdeh, "Navigating the confluence of artificial intelligence and education for sustainable development in the era of industry 4.0: Challenges, opportunities, and ethical dimensions," Journal of Cleaner Production, vol. 437, 2024.
7. A. Ferchaud, J. Grzeslo, S. Orme, and J. LaGroue, "Parasocial attributes and YouTube personalities: Exploring content trends across the most subscribed YouTube channels," Computers in Human Behavior, vol. 80, pp. 88-96, 2018, ISSN: 0747-5632.
8. R. Sajja, Y. Sermet, M. Cikmaz, D. Cwiertny, and I. Demir, "Artificial Intelligence-Enabled Intelligent Assistant for Personalized and Adaptive Learning in Higher Education," submitted to arXiv, Sep. 19, 2023..
7. C. K. Y. Chan and W. Zhou, "Deconstructing Student Perceptions of Generative AI (GenAI) through an Expectancy Value Theory (EVT)-based Instrument," arXiv, 2023.
8. O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education–where are the educators?" International Journal of Educational Technology in Higher Education, vol. 16, no. 1, p. 39, 2019.
9. A. Piktus, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, and D. Kiela, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in Proc. Conf. Neural Information Processing Systems (NeurIPS), Dec. 2020..
10.A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention Is All You Need," arXiv preprint arXiv:1706.03762, submitted Jun. 12, 2017,
11.H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar, A. Rodriguez, A. Joulin, and G. Lample, "LLaMA: Open and Efficient Foundation Language Models,", Feb. 24, 2023. [Online]..
12. H. Naveed, A. U. Khan, S. Qiu, M. Saqib, S. Anwar, M. Usman, N. Akhtar, N. Barnes, and A. Mian, "A Comprehensive Overview of Large Language Models," arXiv preprint arXiv:March 1, 2021
14. M. Lutz, S. Gadaginamath, N. Vairavan and P. Mui, "Examining Political Bias within YouTube Search and Recommendation Algorithms," 2021 IEEE Symposium Series on Computational Intelligence (SSCI), Orlando, FL, USA, 2021, pp. 1-7, doi: 10.1109/SSCI50451.2021.9660012.