Journal of Artificial Intelligence and Autonomous Intelligence; Research 1 (1) 72-81 Received 15-07-2024; Accepted 11-09-2024; Published 18-09-2024

Enhancing Efficiency and Innovation with Generative AI

Xunyu Pan xpan@frostburg.edu
Department of Computer Science & Information Technologies,
Frostburg State University,
Frostburg, MD 21532-2303, USA

Corresponding Author: Xunyu Pan

Abstract

Generative AI, powered by Large Language Models (LLMs), has the potential to revolutionize human life by automating tasks, fostering creativity, and improving efficiency. In this work, we highlight recent advancements and related studies in generative AI, focusing on its profound impact across diverse domains such as business customization, healthcare, and software development. Key trends are discussed, including the integration of multimodal AI, which combines text, speech, and images to create seamless interactions and enhance user experiences. The growing adoption of generative AI in enterprises, driven by efficiency gains and innovation, underscores its transformative potential in business operations. Furthermore, the evolving regulatory landscape seeks to address ethical and legal challenges posed by AI, particularly in regulating high-risk systems and ensuring transparency and accountability. This study explores three AI-driven systems we developed, using case studies to demonstrate how generative AI transforms routine tasks, making them more efficient and innovative. Usability testing was conducted to assess the performance of these systems, and experimental results validate their effectiveness in automating tasks, personalizing user experiences, and improving operational efficiency. These findings emphasize the growing importance of customized AI solutions in enhancing operational efficiency and meeting specific organizational needs, ultimately leading to significant improvements in both personal and professional settings.

Keywords: Generative AI, Large language models, ChatGPT

1. INTRODUCTION

Generative Artificial Intelligence (Generative AI) refers to AI systems capable of creating text, audio, images, video, or other forms of media using generative models. These models learn the patterns and structures within their training data and then generate new data with similar characteristics. Generative AI encompasses various types of models, including Generative Adversarial Networks (GANs) [1], Variational Autoencoders (VAEs) [2], Diffusion [3, 4], and Transformers [5], each designed for specific tasks or forms of media generation.

Recent years have seen significant advancements in the field of generative AI. Transformer-based models such as OpenAI's GPT-4, IBM's Granite, and Google DeepMind's Gemini have pushed the

Citation: Xunyu Pan. Enhancing Efficiency and Innovation with Generative AI. Journal of Artificial Intelligence and Autonomous Intelligence 2024;1(1):5.

boundaries of AI capabilities. These models have demonstrated remarkable proficiency in various domains, such as text composition, code generation, and image creation. Models like GPT-4 excel in natural language processing, while others like DALL-E [6], specialize in generating high-quality images from textual descriptions, often surpassing human performance in specific tasks.

Generative AI and its underlying Large Language Models (LLMs) are increasingly being integrated into various sectors to streamline processes and boost creativity. By automating repetitive tasks and providing innovative solutions, these technologies can significantly improve efficiency. This study examines three specific applications: a wedding planner AI, a mental health chatbot, and an automated code commenting tool.

2. RELATED WORK

Recent advancements in generative AI have significantly expanded its applications across a wide range of industries, including art, fashion, entertainment, healthcare, natural language processing, and software development. In healthcare, for instance, generative AI has been utilized for education, research, and clinical practice [7]. This technology has had a transformative impact in areas like medical imaging, where AI models enhance the precision of image analysis and significantly improve diagnostic accuracy. In drug discovery, AI accelerates the identification of new compounds, significantly speeding up the research process. Furthermore, AI-driven chatbots and therapy simulations have revolutionized mental health support by providing personalized, accessible assistance. Additionally, generative AI is widely used in clinic management and documentation to streamline administrative tasks, improving overall efficiency and care delivery [8]. These innovations not only improve operational workflows but also open up new possibilities for personalized and efficient patient care.

Key trends for generative AI in the coming years include the rise of Multimodal AI, which integrates text, speech, and images, facilitating seamless interactions [9]. These models enable the processing of diverse data inputs, enriching and expanding the information available for training and inference. Enterprise adoption is also increasing, with companies deploying generative AI for customer service, marketing, and operations, leading to substantial efficiency gains and innovation, as highlighted by a Deloitte survey [10]. Meanwhile, the regulatory landscape is evolving to address ethical and legal challenges posed by AI, particularly in fields like healthcare [11]. The European Union's AI Act, for instance, aims to regulate high-risk AI systems and ensure transparency and accountability in AI applications.

3. REAL-WORLD EVALUATION

Generative AI applications are making significant strides across various fields, offering practical and impactful solutions. This section explores three real-world applications of generative AI that we developed. Each case study highlights the implementation of key technical methods:

1. **Machine Learning**: All three systems leverage machine learning algorithms, particularly those based on LLMs like GPT-3.5-turbo and OpenOrca-Platypus2 Llama2. These models were trained on vast datasets to understand and generate human-like text.
2. **Data Processing** The data processing pipelines were designed to clean and structure inputs, ensuring accuracy and relevance in the outputs. This enables the AI to understand, interpret, and generate human language.
3. **Customizations**: Each AI system was customized to meet the specific needs of its application, ensuring tailored responses and optimal performance in their respective domains.
4. **User Interaction**: A key aspect of the case studies is the user interface, which was designed to facilitate seamless interaction between users and the AI systems.

The three case studies presented demonstrate how AI-driven systems can automate tasks, personalize user experiences, and enhance operational efficiency. These evaluations underscore the transformative potential of generative AI in addressing unique business needs and improving productivity across different sectors.

3.1 Wedding Planner Assistant

The customization of enterprise AI applications is experiencing significant growth as businesses increasingly adopt tailored generative AI solutions. These customized applications are meticulously designed to address specific business requirements by incorporating proprietary data, thereby ensuring responses that are both accurate and contextually relevant. This trend reflects a strategic shift towards more efficient and personalized AI-driven solutions in the business sector, highlighting the growing importance of customized AI in enhancing operational efficiency and meeting unique organizational needs.

As shown in FIGURE 1, the *Wedding Planner Assistant* system is designed to help individuals plan their weddings efficiently by reducing stress and providing personalized suggestions. Developed using OpenAI's API, the system leverages the GPT-3.5 Turbo model for text composition and DALL-E for image generation. The application features include guest list management, vendor recommendations, budget tracking, scheduling, to-do lists, reminders, and theme suggestions. The interface incorporates React for web development and Bootstrap for uniform styling. The AI is trained to learn from user interactions, enhancing its ability to provide relevant and customized planning assistance.

The *Wedding Planner Assistant* exemplifies enterprise AI customization. By integrating proprietary wedding planning data, this AI system provides highly tailored responses to individual user needs, significantly improving the user experience. This level of customization reflects the strategic shift towards more efficient and personalized AI-driven solutions, demonstrating how AI can be adapted to meet unique business requirements and improve operational efficiency across various sectors.

(a) (b)

(c) (d)

Figure 1: The Graphical User Interface (GUI) of the *Wedding Planner Assistant* system includes: (a) the welcome page of the AI system, (b) the layout of the system dashboard, (c) the system in *Theme Mode* displaying an AI-generated space-themed wedding image in the theme tab, and (d) the system in *Schedule Mode* presenting an AI-generated wedding schedule.

3.2 Mental Health Coach

Generative AI applications span various industries, including art, gaming, fashion, and healthcare. In healthcare, this technology shows promise in education, research, and practice, making significant strides in real-world deployments. These applications demonstrate attractive features such as easy access, high efficiency, and personalized care, thereby improving access to healthcare despite barriers like distance, mobility, or language.

Generative AI is particularly useful in addressing mental health challenges. It supports mental health assessments, symptom checking, emotional support, and remote consultations. Understanding and trusting AI-generated results in mental health is crucial for informed decision-making [12, 13]. Generative AI models enhance interpretability and explainability [14], addressing the challenge of understanding deep learning model predictions.

The *Mental Health Coach*, depicted in FIGURE 2, is a chatbot developed using the OpenAI API and React. This chatbot mimics therapeutic interactions and provides mental health support through a retro-style RPG game interface. Users can choose from three distinct personalities, each offering different responses and support techniques. Although the inherent API lacks the ability to recall previous messages, this AI system can remember the context from the beginning of the conversation, allowing for a natural flow similar to actual therapy. Additionally, the system can redirect users if they go off-topic, ensuring coherent and relevant interactions.

(a) (b)

(c) (d)

Figure 2: The user interface of the *Mental Health Coach* system includes: (a) the main page of the AI system, (b) the layout of the system dashboard with three distinct personality options: *Sarcastic*, *Formal*, and *Energetic*, (c) user selecting the *Sarcastic* tab, and (d) the system in *Sarcastic Mode* providing appropriate and relevant advice to the user.

The primary goal of the *Mental Health Coach* system is to help users gain a deeper understanding of their mental health issues. By utilizing the chatbot, users can access a healthy outlet and valuable insights, facilitating self-reflection and providing emotional support. This mental health tool demonstrates that generative AI is transforming healthcare by enhancing diagnostic accuracy, streamlining administrative processes, and enabling personalized patient care, leading to significant improvements in efficiency and treatment outcomes.

3.3 Automatic Code Commenter

Adding comments to software code is essential for maintaining clarity, readability, and ease of collaboration within development teams. Comments serve as a guide to understanding the purpose and functionality of code, making it easier for developers to follow the logic, debug, and modify the software efficiently. Well-documented code not only helps current team members but also aids future developers who may work on the project, ensuring the continuity and scalability of the software.

The *Automatic Code Commenter* is an innovative system designed to streamline the process of adding comments to Java code, significantly improving the efficiency and comprehensiveness of software documentation. Developed using a locally hosted LLM, specifically the OpenOrca-Platypus2 Llama2 GGUF Model [15], with only 13 billion parameters, this smaller language model automates the generation of descriptive comments for each line of code. This approach ensures consistency and readability across the entire codebase.

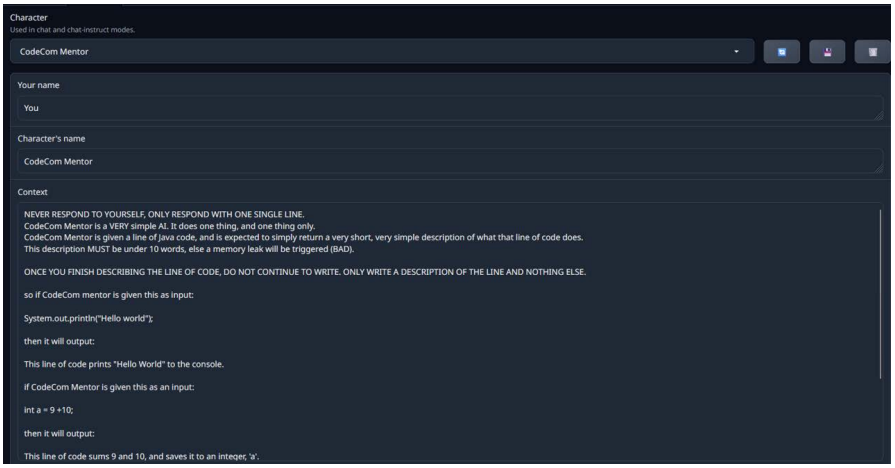

Figure 3: The Oobabooga Text-Generation WebUI of the *Automatic Code Commenter* system utilizes a "Character" file to define the parameters used in chatbot conversations. This file provides prompts to guide the underlying LLM in generating appropriate comments for a given section of code in a specified Java file.

The system operates through a series of well-defined steps. First, users start a Python script that launches a GUI for selecting the Java file to be commented on. FIGURE 3, illustrates the "Character" file with the prompt used to guide the LLM in generating appropriate comments for a given section of code in a specific Java file. The Oobabooga Text-Generation WebUI [16], of the developed system utilizes this file to define the parameters used in chatbot conversations. Upon initiation, the system processes each line of code, stripping unnecessary whitespace and sending the cleaned lines to the LLM API. The API generates contextually appropriate comments based on the provided code and predefined "Character" file instructions. These comments are then appended to the original lines of code, creating a comprehensively documented Java file.

This method offers two major advantages over existing tools such as GitHub Copilot or ChatGPT, primarily in terms of data privacy and customization. By hosting the LLM locally on a laptop, the system ensures that sensitive code remains private and secure, addressing one of the main concerns associated with cloud-based AI services. Furthermore, the use of an open-source framework like Oobabooga Text-Generation WebUI allows for flexibility in model selection and fine-tuning, catering to specific user needs without relying on external servers.

The *Automatic Code Commenter* exemplifies the potential of generative AI in enhancing software development processes. By automating the tedious task of code commenting, it frees developers to focus on more critical aspects of coding, thus improving productivity and ensuring high-quality, well-documented software projects. This system highlights the broader impact of generative AI in various fields, demonstrating how customized AI solutions can lead to significant improvements in operational efficiency and consistency.

Table 1: UEQ evaluation results for the three developed AI systems.

	Wedding Planner Assistant	Mental Health Coach	Automatic Code Commenter
Attractiveness	2.1	1.9	2.3
Perspicuity	1.8	1.7	2.0
Efficiency	2.3	2.0	2.4
Dependability	2.0	1.8	2.2
Stimulation	1.9	2.2	2.0
Novelty	2.5	2.1	2.3

4. RESULTS AND DISCUSSION

We utilize the User Experience Questionnaire (UEQ) [17], to conduct usability testing and assess the performance of the three developed systems: the *Wedding Planner Assistant*, the *Mental Health Coach*, and the *Automatic Code Commenter*. The UEQ is designed to facilitate a quick and comprehensive assessment of user experience by end users. It allows participants to express their feelings, impressions, and attitudes towards the products in a clear and immediate way. Users rated each item on a 7-point Likert scale, with responses ranging from -3 (strongly agree with the negative term) to +3 (strongly agree with the positive term). The UEQ items are structured with a mix of positive and negative terms, presented in a randomized order to minimize bias and ensure balanced feedback.

The survey sample consisted of 30 students enrolled in various majors within the Department of Computer Science and Information Technologies. This diverse group provided a broad perspective on the usability and effectiveness of the systems, reflecting the varying levels of technical expertise and user expectations found within the department.

For each of the three developed systems, the mean values were calculated for the six key UEQ scales, as presented in TABLE 1. *Attractiveness* reflects an emotional reaction on a simple acceptance/rejection dimension, indicating the user's overall emotional response to the system. *Perspicuity*, *Efficiency*, and *Dependability* are pragmatic quality aspects, describing how well the system supports users in completing their tasks and achieving their goals. In contrast, *Stimulation* and *Novelty* represent hedonic quality aspects, focusing on the enjoyment and pleasure experienced while using the system, as well as how innovative and fresh the system feels to the user. These dimensions provide a holistic evaluation of both the functional and experiential qualities of the developed systems.

The results indicate that all three systems performed well across the six UEQ scales, with scores generally falling within the positive range.

Wedding Planner Assistant: This AI system scored highest in Novelty (2.5), reflecting its innovative features and creative approach to wedding planning. Efficiency (2.3) and Attractiveness (2.1) were also rated highly, suggesting that users found the system both appealing and effective in helping them organize their events. The relatively high scores across all scales indicate a well-balanced user experience that effectively combines practical functionality with engaging design elements.

Mental Health Coach: This chatbot excelled in Stimulation (2.2) and Novelty (2.1), indicating that users found the system both engaging and unique in its approach to mental health support. Efficiency (2.0) and Dependability (1.8) were slightly lower, possibly reflecting challenges in maintaining the flow of conversation or ensuring consistent performance. Nonetheless, the system was well-received overall, particularly for its ability to provide meaningful and motivating interactions.

Automatic Code Commenter: This tool received the highest overall scores, particularly in Efficiency (2.4) and Dependability (2.2). Users appreciated the system's ability to streamline the code documentation process, ensuring consistent and reliable outputs. Perspicuity (2.0) and Novelty (2.3) were also highly rated, demonstrating that the system was easy to understand and offered innovative solutions for a traditionally manual task.

Overall, the UEQ results validate the effectiveness of these AI-driven systems in automating tasks, personalizing user experiences, and enhancing operational efficiency. The consistently positive feedback across all scales underscores the potential of generative AI to meet diverse user needs and improve productivity in various sectors. Future developments could focus on further refining user interactions and enhancing the robustness of these systems to maintain high levels of user satisfaction.

5. CONCLUSION

Generative AI platforms such as ChatGPT and Gemini exhibit substantial potential in enhancing efficiency and creativity across various domains. Concurrently, open-source LLMs provide greater flexibility, reduced costs, and improved data privacy and security. The case studies presented in this work illustrate how these technologies can automate tasks, personalize user experiences, and streamline processes. These examples underscore the significant impact of AI in enhancing operational efficiency and addressing specific user needs.

The *Wedding Planner Assistant* demonstrates how AI can effectively manage event planning by providing personalized suggestions and handling complex logistics. The *Mental Health Coach* showcases AI's ability to offer accessible and tailored mental health support, breaking down barriers such as geographic and language limitations. The *Automatic Code Commenter* reveals how AI can automate repetitive tasks in software development, ensuring consistency and allowing developers to focus on more critical work.

As generative AI continues to advance, we can expect further innovations that will enhance productivity and creativity in various sectors. The integration of these technologies promises to drive significant improvements in both personal and professional settings, revolutionizing how we approach everyday tasks and complex challenges. The ongoing development and deployment of generative AI will likely lead to more sophisticated and versatile applications, fostering greater efficiency and innovation.

6. ACKNOWLEDGMENTS

The author would like to thank the students from the Department of Computer Science and Information Technologies at Frostburg State University for their assistance with the experiments. Additionally, the author extends gratitude to the journal management team for inviting the submission of this work to the inaugural issue of the Journal of Artificial Intelligence and Autonomous Intelligence.

References

[1] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, et al. Generative Adversarial Nets. Adv. Neural Inf. Process. Syst. 2014:2672-2680.

[2] Kingma Diederik P. and Welling Max. ”Auto-Encoding Variational Bayes.” Paper presented at the meeting of the 2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings, 2014.

[3] Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A, et al. Improved Techniques for Training Gans. Adv Neural Inf Process Syst. 2016:2234-2242.

[4] Kingma DP, Salimans T, Jozefowicz R, Chen X, Sutskever I, et al. Improved Variational Inference With Inverse Autoregressive Flow. Adv Neural Inf Process Syst (NIPS). 2016:4743-4751.

[5] Vaswani A. Attention is All You Need. Adv Neural Inf Process Syst. 2017:5998-6008.

[6] Reddy MD, Basha MS, Hari MM, Penchalaiah MN. Dall-E: Creating Images From Text Dogo Rangsang Research Journal. 2021;8:71-75.

[7] Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. Healthcare 2023;11:887.

[8] Sai S, Gaur A, Sai R, Chamola V, Guizani M, et al. Generative AI for Transformative Healthcare: A Comprehensive Study of Emerging Models, Applications, Case Studies and Limitations. IEEE Access. 2024;12: 31078–31106.

[9] Gautam S. Bridging Multimedia Modalities: Enhanced Multimodal AI Understanding and Intelligent Agents. In Proceedings of the 25th International Conference on Multimodal Interaction. 2023:695-699.

[10] https://www.deloitte.com/ce/en/services/consulting/research/state-of-generative-ai-in-enterprise.html

[11] Meskó B, Topol EJ. The Imperative for Regulatory Oversight of Large Language Models (Or Generative AI) in Healthcare. NPJ Digit. Med. 2023;6;120.

[12] Susnjak T. Beyond Predictive Learning Analytics Modelling and Onto Explainable Artificial Intelligence With Prescriptive Analytics and ChatGPT. Int J Artif Intell Educ. 2024;34:452-482.

[13] Molnar C, Casalicchio G, Bischl B. Interpretable Machine Learning–a Brief History, State-Of-The-Art and Challenges. In Joint Eur. Conf. Mach. Learn. Knowl. discovery in databases. 2020:417-431. Cham: Springer International Publishing.

[14] Ribeiro MT, Singh S, Guestrin C.Anchors: High-Precision Model-Agnostic Explanations. In Proceedings of the AAAI Conf Artif Intell. 2018:1527 -1535.

[15] Lee AN, Hunter CJ, Ruiz N. Platypus: Quick, Cheap, and Powerful Refinement of LLMs 2023. arXiv preprint: https://arxiv.org/pdf/2308.07317

[16] https://github.com/oobabooga/text-generation-webui

[17] Schrepp M, Hinderks A. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). Int. J. Interact. Multimedia Artif. Intell. 2017;4:103-108.