Comparative Analysis of Generative AI Models

Geeta Rani
CSE APEX
Chandigarh University
Gharuan, Punjab, India
geeta.e15227@cumail.in

Jagjit Singh
CSE APEX
Chandigarh University
Gharuan, Punjab, India
jagjit.e15190@cumail.in

Abhishek Khanna
CSE APEX
Chandigarh University
Gharuan, Punjab, India
abhishek.e13236@cumail.in

2023 International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT) | 979-8-3503-4438-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICAICCIT60255.2023.10465941

***Abstract*—Generative AI models have the potential to create a variety of new content based on training data. They are not only able to create textual content but also other multimedia content such as images, audio, video, etc. They have gained popularity in recent years as they have a major impact on various fields. They are used for several applications from text generation, image generation, and music composition to education, healthcare, and metaverse. Still, several challenges are faced while developing and applying these models i.e., trustworthiness, biased content, overfitting and regulatory concerns. In this paper, the comparative analysis of various generative AI models concerning different parameters is performed with respect to tools, frameworks, input, output, development authority, etc. In addition to these, Applications of different generative AI Models are discussed in various domains**

***Keywords*—Generative AI, GPT, ChatGPT, OpenAI, Generative Adversarial Networks (GANs)**

I. INTRODUCTION

Generative AI models [1] are a fascinating class of artificial intelligence algorithms that can create new and original content. Unlike conventional models that focus on recognizing patterns and making predictions based on existing data, generative AI models can generate entirely new data that is similar to the examples they were trained on.

The potential applications of generative AI are vast, ranging from creative content generation, such as art, music, and storytelling, to practical use cases in various industries, including healthcare, gaming, and design. However, with such creative power comes challenges, such as addressing bias and ethical concerns in the generated content, ensuring the models' outputs are accurate and coherent, and managing the computational resources required for training and inference.

As research in generative AI continues to advance, these models have the potential to revolutionize the way we interact with technology, paving the way for more interactive and personalized experiences in the future. Nevertheless, striking a balance between innovation and responsibility remains crucial in harnessing the full potential of generative AI while mitigating any associated risks

II. GENERATIVE AI VS. TRADITIONAL AI MODELS

Generative AI is used to generate multimedia content i.e. audio, video, images, etc. On the other hand, Traditional AI Models are used for prediction, classification, and regression using decision trees or Support Vector Machine. Generative AI models might face challenges in capturing complex relationships in tabular data, while traditional AI models can encounter issues like overfitting and handling non-linear relationships. Generative AI has become more popular these days as it generates creative content but requires a huge amount of data for effective learning. A broad comparison is shown in Table 1.1.

TABLEI
COMPARISON OF GENERATIVE AI MODELS AND TRADITIONAL MODELS

Aspect	Generative AI Models	Traditional AI Models
Purpose	Primarily used for data generation and synthesis.	Mainly employed for data analysis, forecasting, and classification.
Learning Approach	Unsupervised or semi-supervised learning.	Supervised, unsupervised, or semi-supervised learning.
Model Architecture	Complex neural networks, e.g., Variational Autoencoders (VAEs).	Diverse algorithms like decision trees, regression techniques, SVMs, etc.
Applications	Used in creative tasks (e.g., image, music, text generation).	Applied in various domains (finance, healthcare, business) for data analysis.
Data Requirements	May require larger amounts of data for effective learning.	Can perform better with medium-sized datasets.
Focus on Data Generation	Yes	No
Focus on Data Analysis	No	Yes
Challenges	Capturing complex relationships in tabular data.	Overfitting and handling non-linear relationships.

III. DIFFERENT GENERATIVE AI MODELS

There are many generative models that can create sophisticated content using machine learning and deep learning techniques. Some notable generative AI models are Generative Adversarial Networks (GANs), Generative Pre-trained Transformers (GPT), Variational Autoencoders (VAEs), Autoregressive Models, Flow-based Models, and Style Transfer Networks.

- Generative Adversarial Networks (GANs) [3] work on two neural networks, a generator, and a discriminator. They are trained collectively in a competitive setting. The generator aims to produce realistic data to deceive the discriminator. The discriminator attempts to discriminate between actual and produced data. Over time, GANs become adept at generating increasingly realistic content.

- The Generative Pre-Trained Transformer (GPT) is an advanced NLP model family created by OpenAI. It is created on the Transformer architecture. It undergoes pre-training on extensive internet text data to learn language patterns and grammar. GPT can produce human-like text as a generative model when given a starting prompt. GPT-3, the largest version with 175 billion parameters, exhibits exceptional language capabilities and surpasses other models in various language tasks. Notably, it can perform few-shot and zero-shot learning, requiring minimal examples or no training data for new tasks. GPT's API accessibility enables seamless integration into applications for language understanding and generation, revolutionizing natural language processing.

- Variational Autoencoders (VAEs) are generative models that use an encoder-decoder architecture. They map input data to a continuous latent space and introduce stochasticity by sampling from a distribution. VAEs are trained to optimize two objectives: a reconstruction loss and a regularization term that encourages a structured latent space. The learned latent space enables VAEs to generate new data points by sampling and decoding.

- Autoregressive models are statistical models or neural network architectures that generate sequential data one element at a time, conditioned on previous elements. They model the conditional probability of each element in the sequence based on its preceding elements. Commonly used in tasks like time series forecasting, language modeling, and music generation, autoregressive models are trained using maximum likelihood

- estimation and have applications in various domains that involve sequential data analysis.

- Flow-based models are generative models that use invertible transformations to map data from a simple distribution to a complex one, allowing for efficient sampling and likelihood estimation. They involve latent variables, and the likelihood of data points is computed by integrating these variables. Flow-based models are trained using maximum likelihood estimation and have applications in data density estimation, image generation, etc.

- Style transfer networks are deep learning architectures that blend the content of one image with the artistic style of another image. They utilize pre-trained convolutional neural networks to extract content and style representations from input images. The models optimize a loss function that includes content and style losses to generate the stylized output. Style transfer networks find applications in artistic image editing and image-to-image translation tasks, allowing users to create visually appealing and unique artistic effects.

IV. Generative AI Models Based on Input and Output

Generative AI Models are not able to produce only text but also other multimedia forms i.e., Image, Video, Audio, code, etc. They are subdivided into six categories as shown in fig. 1. depending upon different input and output multimedia content.

A) Text to Text: Models like ChatGPT3 [1], LaMDA [4], and PEER [5] and speech from the brain [6] based models generate human-like text based on a given text prompt

B)Text to Image: These text-to-image generative AI models can generate realistic and diverse images based on textual input, making them valuable tools for creative content generation, art, and multimedia applications. Common examples are DALL-E2 [7], Stable Fusion [8], Imagen [9], and Muse [10].

C)Text to 3D: They are comparatively complex as thequality of results heavily depends on the model architecture, training data, and the richness of the 3D shape representations. Examples are Dream Fusion [11], Magic3D [12].

*D)Image to Text:*There were several state-of-the-art image-to-text generative AI models. These models belong to a category of neural networks known as image captioning models. They take an input image and generate a descriptive textual caption that represents the content of the image. Popular examples are Flamingo [13] and Visual GPT [14]. CLIP, DenseCap.

*E) Text to Video:*There is no specific model solely for this purpose. Yet there are certain models based upon existing architecture these are Phenaki [15], Sounify [16], VideoGPT, etc.

*F) Text to Audio:*These models are designed to convert text input into natural-sounding speech. A few examples are AudioLM [17], Whisper [18], and Jukebox [19].

G) Text-to-Code: Text-to-code generation is a challenging task that involves converting natural language descriptions or instructions into executable code. They may not always produce perfect or optimized code**.** So, it needs to be verified carefully before using it. Common examples are OpenAI Codex [20], Alphacode [21], CodeBert, etc.

H) Text to Science: These models can be powerful tools for generating scientific content, they should be used with caution. The generated text should always be thoroughly reviewed and validated by domain experts before being considered scientifically accurate and reliable. Examples are Galactica, Minerva, etc.

TABLE II
COMPARISON OF DIFFERENT GENERATIVE AI MODELS

Model	Description	Main Focus	Examples of Applications	Tools/Frameworks	Advantages	Disadvantages
Generative Adversarial Networks (GANs)	Consists of a generator and discriminator.	Data generation and realism	Art generation, image synthesis, data augmentation	NVIDIA StyleGAN2, PyTorch-GAN, TensorFlow-GAN	Produces high-quality and diverse outputs	Can be challenging to stabilize and may suffer from mode collapse
Generative Pre-trained Transformer (GPT)	Based on the Transformer architecture.	Text generation	Natural language processing, chatbots, text generation	OpenAI GPT-3, GPT-2, Hugging Face Transformers	Produces coherent and contextually relevant text	May produce plausible but incorrect or nonsensical responses
Variational Autoencoders (VAEs)	Utilizes an encoder and decoder.	Data generation and controllable synthesis	Image generation, data compression, anomaly detection	Keras, TensorFlow Probability, PyTorch	Allows for controlled data generation and manipulation	Output quality may not be as high as GANs
Autoregressive Models	Generates data sequentially based on conditional probabilities.	Text and sequential data generation	Text-to-speech systems, language modeling	OpenAI GPT-2, Google Magenta, Fastai	Provides exact likelihood of data	A slower generation process with, sequential nature can limit parallelism
Flow-based Models	Learns a mapping using invertible transformations.	Data generation and density estimation	Image synthesis, data augmentation	Glow, NICE, Pyro, TensorFlow Probability	Efficient sampling and likelihood calculation	Can be computationally expensive for complex models
Style Transfer Networks	Manipulates the style of an input while preserving its content.	Artistic style transfer	Artistic photo filters, image-to-image translation	Prisma, Neural Style Transfer, Fast Neural Style Transfer	Allows for artistic manipulation of images	May not handle complex images or semantics as effectively

Fig 1: Comparison of Generative Models based on Input and Output

V. GENERATIVE AI MODELS BASED ON THE DEVELOPMENT AUTHORITY

Fig.2. shows the classification as per the developing companies. It has been seen that six development companies are major players in developing these models.

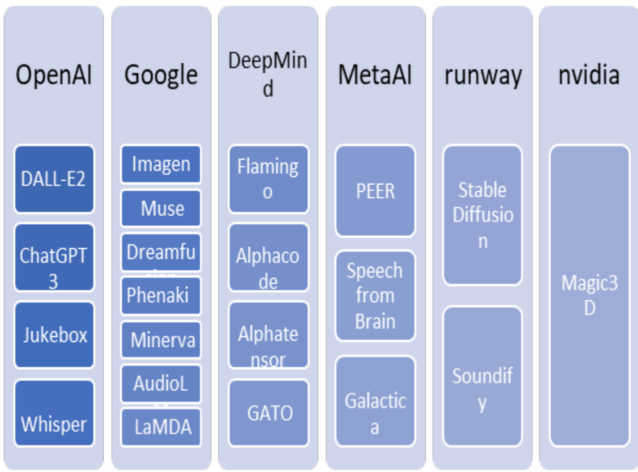

Fig 2: Comparison of Generative Models based on the Developer

VI. APPLICATIONS

Generative AI models have found applications in various domains due to their ability to create new content, simulate data, and generate contextually relevant outputs. Here are some specific applications across different industries

A)Education: Generative AI models offer transformative applications in the field of education [22], revolutionizing how students learn and interact with educational content. One prominent application is content generation, where AI can create personalized study materials catering to each student. Furthermore, AI-powered virtual instructors can provide personalized guidance and support, answering students' questions and adapting to their individual progress. Another vital aspect is language translation, allowing educational resources to be easily accessible to students worldwide, regardless of their native language. Furthermore, AI assists students in improving their writing skills by offering grammar corrections, writing suggestions, and valuable feedback on essays and compositions. Lastly, generative AI models can create interactive simulations and visualizations to help students grasp complex concepts more effectively, making learning more engaging and effective. These applications have the potential to transform traditional education methods and empower both students and educators to achieve better learning outcomes.

*B)Drug Discovery:*Generative AI models have emerged as powerful tools in drug discovery [23], significantly accelerating the process of identifying potential new drug candidates. These models aid in virtual screening and lead optimization, reducing the time and cost required for traditional experimental approaches. Moreover, generative AI can explore vast chemical spaces, uncovering novel compounds that may not have been explored otherwise. This technology plays a crucial role in predicting molecular properties, such as bioactivity and toxicity, guiding researchers in selecting promising drug candidates for further development and clinical trials. Ultimately, the applications of generative AI in drug discovery hold tremendous promise for revolutionizing the pharmaceutical industry and advancing the discovery of new life-saving medications.

C) Medical: Generative AI models have numerous applications in the medical field, transforming various aspects of healthcare [24]. One significant application is medical image generation, where AI can synthesize high-quality medical images to aid in training and augmenting datasets for diagnostic algorithms. In personalized medicine, generative AI can create tailored treatment plans based on patient's genetic profiles and medical histories. Additionally, AI-powered virtual assistants can assist healthcare professionals in diagnosing diseases and recommending appropriate treatments, improving patient care and outcomes, and plastic surgeries [25]. Overall, generative AI models have the potential to revolutionize medical research, diagnosis, and treatment, advancing the field of healthcare to new heights.

D) Gaming and Entertainment: Generative AI has revolutionized the gaming [26] and entertainment industries, opening up new creative possibilities and enhancing user experiences. One prominent application is procedural content generation, where AI models create vast and diverse game environments, landscapes, and levels, ensuring endless and dynamic gameplay experiences for players. Additionally, generative AI is utilized in character and creature design, generating unique and imaginative entities that enrich game worlds. These models also play a crucial role in storytelling and narrative generation, enabling interactive and branching storylines that respond to player choices, immersing them in engaging and personalized gaming adventures. Moreover, AI-powered tools produce stunning visual assets, animations, and sound effects, elevating the overall aesthetics and audio experience of games. Whether in virtual reality or augmented reality, generative AI creates realistic and interactive worlds, providing players with immersive and captivating experiences. As technology advances, generative AI continues to push the boundaries of game design, user engagement, and entertainment, shaping the future of gaming and interactive media.

E) Virtual Reality and Augmented Reality: Generative AI applications in Virtual Reality (VR) and Augmented Reality (AR) have revolutionized the immersive experience, offering users interactive and lifelike encounters. One key application lies in content creation, where AI models generate realistic 3D assets, environments, and characters, speeding up development and enhancing visual quality. Scene reconstruction enables seamless integration of physical spaces into AR experiences, while gesture and pose recognition empower users to interact naturally with virtual objects and characters. Generative AI facilitates interactive storytelling by dynamically generating narratives that respond to user actions, providing personalized and engaging journeys. Additionally, AI-driven simulations simulate real-world scenarios, contributing to authentic and educational experiences in both VR and AR. These advancements in generative AI redefine the possibilities of interactive entertainment, training, marketing, and education, creating a new era of immersive and captivating virtual and augmented experiences.

VII. CHALLENGES AND ETHICS

Generative AI poses several challenges that need to be navigated for its widespread and responsible adoption. One significant challenge is data bias and fairness, as generative AI models can perpetuate biases present in the training data, leading to unfair or discriminatory outputs. Ethical use and potential misuse of generative AI raise concerns about the creation of deepfakes, misinformation, and fake content, which can have harmful implications. The lack of transparency and explainability in complex generative AI models poses difficulties in understanding their decision-making process, hindering trust and accountability. Additionally, ensuring data privacy and security while handling large datasets required for training generative AI is crucial to prevent data breaches and misuse of sensitive information. Addressing these challenges will require collaborative efforts to develop robust regulations, ethical guidelines, and transparency practices, ensuring that generative AI technology is harnessed responsibly for the benefit of society.

Ethics and regulations in Generative AI are critical to addressing the challenges [27] and potential risks associated with this transformative technology. From an ethical standpoint, it is essential to consider issues like bias and fairness, ensuring that generative AI models do not perpetuate harmful stereotypes or discriminate against specific groups. Privacy and data protection are crucial concerns, requiring vigorous safety measures to protect sensitive information used in training these models. The creation and dissemination of fake content and deepfakes raise ethical questions about the responsible use of generative AI. From a regulatory perspective, clear guidelines are required to govern the development, deployment, and monitoring of generative AI applications. Regulations should address issues like data privacy, security, and transparency, and establish responsible practices [28]. Collaborative efforts are essential to strike a balance between fostering innovation and ensuring that Generative AI is harnessed responsibly, adhering to ethical principles and regulatory frameworks that protect individuals and society as a whole

VIII. CONCLUSIONS

Generative AI models have revolutionized various domains by enabling machines to create content that is both creative and contextually relevant. These models have demonstrated impressive capabilities in tasks like image generation, text generation, code synthesis, and even scientific content creation. They leverage advanced techniques such as neural networks, transformers, and attention mechanisms to learn patterns and relationships within data to produce new content.

Generative AI models have significant potential to enhance human creativity, automate content generation, and assist in various tasks that require creative outputs. They have shown promise in art, design, storytelling, language translation, and even scientific research. Moreover, these models have been adapted for personalized experiences, enabling fine-tuning and transfer learning to align the generated content with individual preferences.

However, it is crucial to remain aware of the ethical implications and challenges associated with generative AI models. Concerns regarding data privacy, bias, misuse, and the potential for generating fake content must be addressed to ensure responsible deployment and usage of these models.

As AI research continues to progress, generative AI models are expected to become even more sophisticated and capable, further empowering industries and enriching human experiences. By combining cutting-edge technology with responsible practices, the potential of generative AI to drive positive advancements and innovations across various domains can be harnessed.

REFERENCES

[1] R. Gozalo-Brizuela and E. C. Garrido-Merchan, “ChatGPT is not all you need. A State of the Art Review of large Generative AI models,” arXiv:2301.04655 [cs], Jan. 2023, Available: https://arxiv.org/abs/2301.04655

[2] R. Salakhutdinov, “Learning Deep Generative Models,” Annual Review of Statistics and Its Application, vol. 2, no. 1, pp. 361–385, Apr. 2015, doi: https://doi.org/10.1146/annurev-statistics-010814-020120..

[3] D. Saxena and J. Cao, “Generative Adversarial Networks (GANs): Challenges, Solutions, and Future Directions,” arXiv:2005.00065 [cs, eess, stat], May 2020, Available: https://arxiv.org/abs/2005.00065

[4] R. Thoppilan, “LAMDA: Language Models for Dialog Applications,” arXiv.org, Jan. 20, 2022. https://arxiv.org/abs/2201.08239

[5] T. Schick et al., “PEER: A Collaborative Language Model,” arXiv.org, Aug. 24, 2022. https://arxiv.org/abs/2208.11663 (accessed Sep. 29, 2023).

[6] A. Défossez, C. Caucheteux, J. Rapin, O. Kabeli, and J.-R. King, “Decoding speech from non-invasive brain recordings,” arXiv.org, Aug. 25, 2022. https://arxiv.org/abs/2208.12266

[7] G. Daras and A. G. Dimakis, “Discovering the Hidden Vocabulary of DALLE-2,” arXiv:2206.00169 [cs], May 2022, Available: https://arxiv.org/abs/2206.00169

[8] C. Saharia et al., “Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding,” openreview.net, Oct. 31, 2022. https://openreview.net/forum?id=08Yk-n5l2Al

[9] D. P. Kingma, T. Salimans, B. Poole, and J. C. Ho, “Variational Diffusion Models,” Jul. 2021, doi: https://doi.org/10.48550/arxiv.2107.00630.

[10] H. Chang et al., “Muse: Text-To-Image Generation via Masked Generative Transformers,” arXiv.org, Jan. 02, 2023. https://arxiv.org/abs/2301.00704

[11] B. Poole, A. Jain, J. T. Barron, and B. Mildenhall, “DreamFusion: Text-to-3D using 2D Diffusion,” arXiv:2209.14988 [cs, stat], Sep. 2022, Available: https://arxiv.org/abs/2209.14988

[12] “Magic3D: High-Resolution Text-to-3D Content Creation,” research.nvidia.com. https://research.nvidia.com/labs/dir/magic3d/ (accessed Sep. 29, 2023).

[13] J.-B. Alayrac et al., “Flamingo: a Visual Language Model for Few-Shot Learning,” arXiv:2204.14198 [cs], Nov. 2022, Available: https://arxiv.org/abs/2204.14198

[14] P. Budzianowski and I. Vulić, “Hello, It’s GPT-2 -- How Can I Help You? Towards the Use of Pretrained Language Models for Task-Oriented Dialogue Systems,” arXiv:1907.05774 [cs], Aug. 2019, Available: https://arxiv.org/abs/1907.05774

[15] R. Villegas et al., “Phenaki: Variable Length Video Generation From Open Domain Textual Description,” arXiv.org, Oct. 05, 2022. https://arxiv.org/abs/2210.02399 (accessed Sep. 29, 2023)

[16] David Chuan-En Lin, A. Germanidis, C. Valenzuela, Y.-N. Shi, and Nikolas Martelaro, “Soundify: Matching Sound Effects to Video,” Dec. 2021, doi: https://doi.org/10.48550/arxiv.2112.09726

[17] Z. Borsos et al., “AudioLM: a Language Modeling Approach to Audio Generation,” arXiv (Cornell University), Sep. 2022, doi: https://doi.org/10.48550/arxiv.2209.03143.

[18] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, “Robust Speech Recognition via Large-Scale Weak Supervision,” arXiv:2212.04356 [cs, eess], Dec. 2022, Available: https://arxiv.org/abs/2212.04356

[19] P. Dhariwal, H. Jun, C. Payne, J. W. Kim, A. Radford, and I. Sutskever, “Jukebox: A Generative Model for Music,” arXiv:2005.00341 [cs, eess, stat], Apr. 2020, Available: https://arxiv.org/abs/2005.00341

[20] M. Chen et al., “Evaluating Large Language Models Trained on Code,” arXiv:2107.03374 [cs], Jul. 2021, Available: https://arxiv.org/abs/2107.03374

[21] Y. Li et al., “Competition-level code generation with AlphaCode,” Science, vol. 378, no. 6624, pp. 1092–1097, Dec. 2022, doi: https://doi.org/10.1126/science.abq1158.

[22] D. Baidoo-Anu and L. Owusu Ansah, “Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning,”

papers.ssrn.com, Jan. 25, 2023. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4337484

[23] B. Tang, J. Ewalt, and H.-L. Ng, "Generative AI Models for Drug Discovery," Biophysical and Computational Tools in Drug Discovery, pp. 221–243, 2021, doi: https://doi.org/10.1007/7355_2021_124.

[24] A. Kleebayoon and V. Wiwanitkit, "ChatGPT in medical practice, education and research: malpractice and plagiarism," Clinical Medicine (London, England), vol. 23, no. 3, p. 280, May 2023, doi: https://doi.org/10.7861/clinmed.Let.23.3.2.

[25] R. Gupta, I. Herzog, J. Weisberger, J. Chao, K. Chaiyasate, and E. S. Lee, "Utilization of ChatGPT for Plastic Surgery Research: Friend or Foe?," Journal of Plastic, Reconstructive & Aesthetic Surgery, Mar. 2023, doi: https://doi.org/10.1016/j.bjps.2023.03.004.

[26] J. Ratican, J. Hutson, and A. Wright, "A Proposed Meta-Reality Immersive Development Pipeline: Generative AI Models and Extended Reality (XR) Content for the Metaverse," Journal of Intelligent Learning Systems and Applications, vol. 15, no. 01, pp. 24–35, 2023, doi:

https://doi.org/10.4236/jilsa.2023.151002.

[27] Fiona Fui-Hoon Nah, R. Zheng, J. Cai, K. Siau, and L. Chen, "Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration," Journal of information technology case and application research, vol. 25, no. 3, pp. 1–28, Jul. 2023, doi:

https://doi.org/10.1080/15228053.2023.2233814.

[28] Fiona Fui-Hoon Nah, R. Zheng, J. Cai, K. Siau, and L. Chen, "Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration," Journal of information technology case and application research, vol. 25, no. 3, pp. 1–28, Jul. 2023, doi:

https://doi.org/10.1080/15228053.2023.2233814.