

## Chapter 5

# Role of machine learning and deep learning in advancing generative artificial intelligence such as ChatGPT

Nitin Liladhar Rane <sup>1</sup>, Suraj Kumar Mallick <sup>2</sup>, Ömer Kaya <sup>3</sup>, Jayesh Rane <sup>4</sup>

<sup>1</sup> Vivekanand Education Society's College of Architecture (VESCOA), Mumbai 400074, India

<sup>2</sup> Shaheed Bhagat Singh College, University of Delhi, New Delhi 110017, India

<sup>3</sup> Engineering and Architecture Faculty, Erzurum Technical University, Erzurum 25050, Turkey

<sup>4</sup> Pillai HOC College of Engineering and Technology, Rasayani, India

[nitinrane33@gmail.com](mailto:nitinrane33@gmail.com)

**Abstract:** The advancement of machine learning (ML) and deep learning (DL) has greatly accelerated the progress of generative artificial intelligence (GAI) models such as ChatGPT, transforming multiple industries through improved human-machine communication. This study investigates how ML and DL are crucial for the development of GAI, with a specific emphasis on their architectures, methods, and uses that have propelled its advancement. Cutting-edge models, especially transformer-based designs, have shown remarkable abilities in natural language processing (NLP), allowing for the creation of coherent, contextually appropriate, and human-like text. The combination of large quantities of data and advanced algorithms like reinforcement learning and unsupervised learning has improved these models, making them better at comprehending and producing language with incredible precision. Further, the progress in computer speed and the access to vast amounts of data have accelerated the development of GAI, enabling the training of models with billions of parameters. This study outlines the various ways ChatGPT can be used in customer service, content creation, and education, underscoring its ability to enhance human productivity and creativity. It also focuses on the ethical aspects and difficulties related to GAI, such as reducing bias, ensuring transparency, and responsibly deploying AI. This research offers a thorough examination of how ML and DL are influencing generative AI's future through analyzing recent trends and advancements, leading to the development of smarter and more interactive systems.

**Keywords:** Artificial intelligence, Machine learning, Deep learning, Generative artificial intelligence, ChatGPT, Natural language processing.

---

**Citation:** Rane, N. L., Mallick, S. K., Kaya, O., & Rane, J. (2024). Role of machine learning and deep learning in advancing generative artificial intelligence such as ChatGPT. In *Applied Machine Learning and Deep Learning: Architectures and Techniques* (pp. 96-111). Deep Science Publishing. [https://doi.org/10.70593/978-81-981271-4-3\\_5](https://doi.org/10.70593/978-81-981271-4-3_5)

---

## 5.1 Introduction

The fast development of artificial intelligence (AI) has caused significant changes in many areas, with generative AI being highlighted as a revolutionary advancement (Minaee et al., 2024; Raiaan et al., 2024; Chang et al., 2024). Generative AI, like ChatGPT, has shown remarkable abilities in both understanding and producing natural language (Raiaan et al., 2024; Hadi et al., 2023; Kukreja et al., 2024; Hadi et al., 2023). These advanced models use advanced machine learning (ML) and deep learning (DL) methods to generate text that resembles human writing, allowing for various applications from automated customer support to creating creative content (Raiaan et al., 2024; Kaur et al., 2024; Jovanovic & Voss, 2024; Wang et al., 2024). The foundations of generative AI are built upon machine learning and deep learning (Raiaan et al., 2024; Bharathi Mohan et al., 2024; Yan et al., 2024; Myers et al., 2024). These technologies make it easier to handle and examine large quantities of data, enabling AI models to identify patterns and produce coherent and contextually appropriate results. Advancements in generative AI have greatly improved due to the shift from rule-based systems to advanced neural networks (Hadi et al., 2023; Veres, 2022; Lam et al., 2024; Thirunavukarasu et al., 2023). This research examines how ML and DL contribute to the progress of generative AI, specifically highlighting the impact on models such as ChatGPT that have shown outstanding performance.

Incorporating ML and DL into generative AI requires various essential elements (Hadi et al., 2023; Thirunavukarasu et al., 2023; Su et al., 2024). Deep learning, employing hierarchical neural networks, enables the extraction of intricate features from data, resulting in enhanced and more detailed text creation (Raiaan et al., 2024; Pahune & Chandrasekharan, 2023; Xu et al., 2024; Karanikolas et al., 2023). Methods like transformers have enhanced the field even more by allowing models to manage long-distance connections and create suitable responses in context (Hadi et al., 2023; Alwahedi et al., 2024; Zhao et al., 2024; Shi et al., 2024). Moreover, improvements in hardware and computational capabilities have increased the speed of training and implementation of these advanced models, making them usable and feasible for real-world uses (Hadi et al., 2023; Yue et al., 2023; Min et al., 2023; Cheng, 2024).

Table 5.1 summarizes the role of various machine learning (ML) and deep learning (DL) techniques used in the development of generative AI models such as ChatGPT. Each technique explains its contribution and benefits to generative AI. These techniques provide a wide range of benefits, from increasing language understanding and production capabilities to optimizing model performance (Hsu & Ching, 2023; Cantens, 2024; Bridges et al., 2024; Kalota, 2024). In this way, generative AI models can produce more accurate, contextual and believable texts.

**Table 5.1 ML and DL Techniques for Generative Artificial Intelligence**

<b>Technical</b>	<b>Description</b>	<b>Contribution to Generative AI</b>
Natural Language Processing (NLP)	Techniques used to understand and produce human language.	It increases the language understanding and production capabilities of models such as ChatGPT.
Recurrent Neural Networks (RNN)	Neural networks used for time series and sequential data processing.	It helps language models better understand sentence structures and contexts.
Transformer Models	Advanced neural networks developed specifically for language processing.	It enables ChatGPT to generate more meaningful and contextual responses by using attention mechanisms.
Large Language Models (LLM)	Language models trained on very large datasets.	It allows them to have a broader knowledge base and provide more accurate answers.
Transfer Learning	Adapting pre-trained models for new tasks.	It allows models such as ChatGPT to quickly adapt to different areas and tasks.
Generative Adversarial Networks (GANs)	Two neural networks compete with each other to produce more realistic data.	It helps models like ChatGPT produce more realistic and believable texts.
Hyperparameter Optimization	Determining the most appropriate parameters to increase the performance of the models.	Improves ChatGPT's response quality and accuracy.
Fine-tuning	Fine-tuning a model for a specific task or dataset.	It allows ChatGPT to better adapt to specific use cases.
Big Data Analytic	Processing and analyzing very large data sets.	It allows ChatGPT to access a broader knowledge base and use that knowledge in a meaningful way.

The main goal of this research is to give a detailed overview of the current status of machine learning and deep learning in enhancing generative AI, specifically focusing on ChatGPT. This research adds to our understanding of how advancements in AI-driven text generation are influencing the future by exploring the underlying technologies, important developments, and future possibilities.

#### Significance of the research study

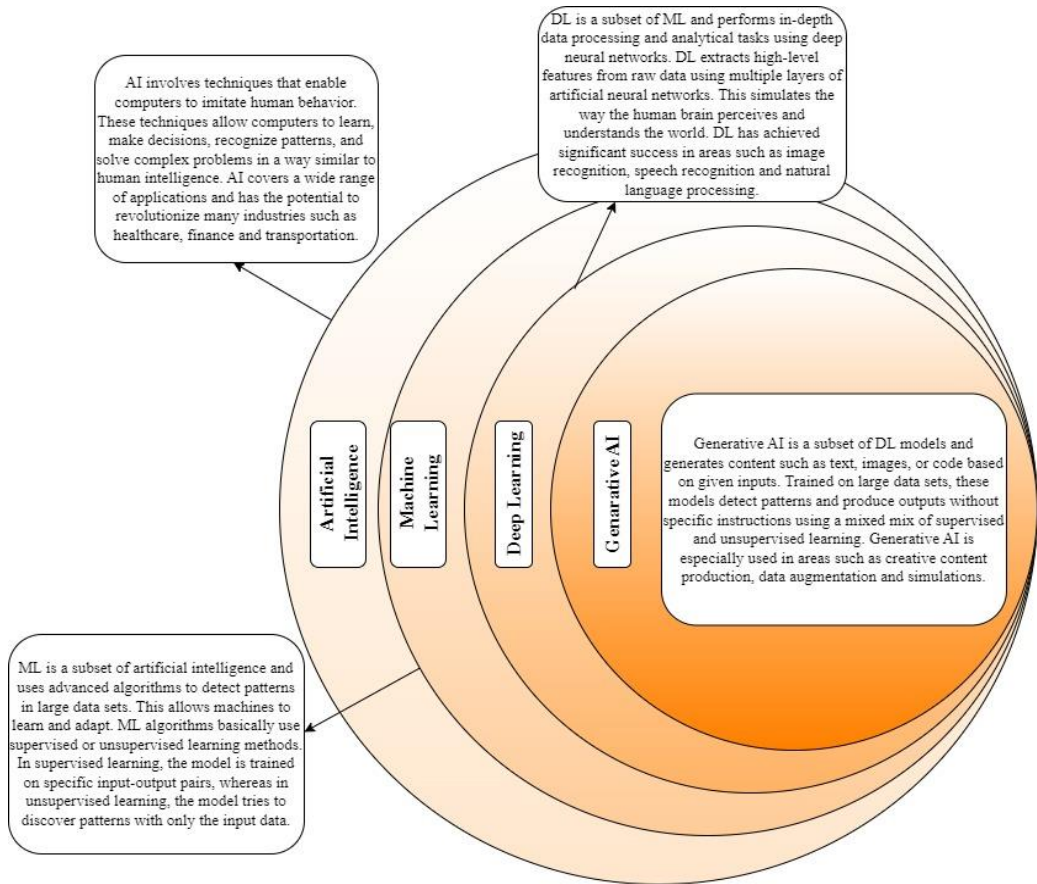
- 1) This research offers an in-depth review of the recent progress in machine learning and deep learning technologies that support generative AI models such as ChatGPT.

- 2) It provides a closer look at the ML and DL methods that have greatly influenced the advancement and efficiency of generative AI, emphasizing important innovations and approaches.
- 3) The study highlights upcoming patterns and possible future paths in the industry, offering a guide for additional progress and research prospects in generative AI.

## 5.2 Methodology

The literature search plan was created to encompass a wide variety of studies related to machine learning (ML), deep learning (DL), and generative artificial intelligence (GAI), specifically highlighting models such as ChatGPT. An extensive search was carried out across various academic databases such as Google Scholar, IEEE Xplore, ScienceDirect, and PubMed. The search utilized keywords such as "machine learning," "deep learning," "generative artificial intelligence," "ChatGPT," "transformer models," "natural language processing," "AI in NLP," and "language generation models." Criteria for including or excluding literature were created in order to ensure the appropriateness and excellence of the chosen material. Inclusion criteria encompassed studies published in peer-reviewed journals or conferences, centered on ML and DL techniques for generative AI, and addressing advancements, applications, or challenges of models such as ChatGPT within the past ten years. Research that was not peer-reviewed, mainly concentrated on unrelated subjects, or released before 2013 unless considered groundbreaking works were eliminated from the studies. The process of data extraction consisted of methodically gathering details from every chosen study, such as authors, publication year, research goals, methodologies, main discoveries, and impact on the field. A thematic analysis was carried out to find common themes, trends, and patterns in the literature, with a focus on the impact of ML and DL on the improvement and creation of generative AI models, specifically ChatGPT.

Fig. 5.1 highlights the interrelationship of these technologies and their importance in the development of generative artificial intelligence, especially models such as ChatGPT. This evolution from artificial intelligence to generative AI contributes to the development of AI systems that enable more sophisticated and human-like interactions (Zhuhadar & Lytras, 2023). Generative AI, as an extension of deep learning models, generates content such as text, images or code. These models are trained on large data sets and create creative and believable content based on the inputs.



**Fig 5.1** The Role of ML and DL in the Development of Generative AI

## 5.3 Results and discussions

### Role of ML in advancing generative artificial intelligence such as ChatGPT

ML has significantly boosted the progress of generative artificial intelligence (AI), especially in creating advanced models such as ChatGPT (Raiaan et al., 2024; Chang et al., 2024; Lappin, 2024; Liu et al., 2024; Park et al., 2024). This advancement represents a major stride in AI abilities, allowing machines to produce text resembling that of humans, tackle intricate language challenges, and engage with users in more organic and instinctive manners (Hadi et al., 2023; Wei et al., 2023; Xie et al., 2023; Yang et al., 2023). Machine learning plays a complex role in this situation, including progress in algorithms, data handling, model structure, and practical implementations. One major way machine learning has boosted generative AI is by creating sophisticated neural network structures (Raiaan et al., 2024; Xu et al., 2024; Hajikhani & Cole, 2024; Omiye et al.,

2024; Kalyan, 2023). Traditional models were constrained because they couldn't handle and produce intricate language patterns. Nevertheless, the field has been transformed with the emergence of deep learning, especially with advancements in architectures such as transformers. Transformers make use of self-attention mechanisms to enhance the capability of models in evaluating the significance of various words within a sentence. This advancement has allowed models such as ChatGPT to produce logical and contextually fitting answers, as it is better at capturing long-distance connections in text compared to older models (Raiaan et al., 2024; Chang et al., 2024; Hadi et al., 2023; Thirunavukarasu et al., 2023; Thirunavukarasu et al., 2023; Su et al., 2024).

The enormous amount of training data and computational resources are crucial for driving the advancement of generative AI through machine learning. The reason for the success of models such as ChatGPT lies in the extensive text data they are trained on, covering a wide range of topics and languages. This thorough training enables the models to comprehend and produce text that is both contextually pertinent and stylistically varied. The ability to access extensive datasets, along with the use of strong computational resources like GPUs and TPUs, has enabled the training of these large models. This ability to utilize large data and computational abilities has been crucial in advancing the potential of generative AI. Machine learning is important for enhancing generative models through a process of continual refinement. Methods like transfer learning and fine-tuning have played a crucial role. Transfer learning enables a model that has been trained on a vast dataset to be customized for particular tasks using smaller datasets, greatly enhancing effectiveness. Fine-tuning increases the model's effectiveness by educating it with data specific to a particular domain, improving its performance in specific situations. This strategy is clearly seen in the implementation of ChatGPT, where the fundamental model is adjusted with an emphasis on conversational information, allowing it to produce responses that are more pertinent and contextually aware. Table 5.2 shows the role of ML and DL in advancing generative artificial intelligence such as ChatGPT

**Table 5.2** Role of ML and DL in advancing generative artificial intelligence such as ChatGPT

References	Aspect	ML	DL	Impact on Generative AI (e.g., ChatGPT)
(Raiaan et al., 2024; Chang et al., 2024; Hadi et al., 2023; Su et al., 2024; Pahune &	Foundational Framework	ML traditionally underpins predictive analytics and classification	DL notably leverages neural networks with multifaceted layers,	This undergirds generative AI models, such as ChatGPT, by furnishing

Chandrasekharan, 2023)		tasks through algorithmic paradigms.	engendering intricate data representations.	fundamental structural scaffolding.
(Hadi et al., 2023; Jovanovic & Voss, 2024; Thirunavukarasu et al., 2023; Thirunavukarasu et al., 2023)	Data Manipulation	ML necessitates meticulous feature engineering and preprocessing methodologies for optimal performance.	DL autonomously abstracts features from raw data, circumventing the need for extensive preprocessing.	This expedites the handling of voluminous datasets with minimal preprocessing overhead, fortifying model agility.
(Lam et al., 2024; Thirunavukarasu et al., 2023; Liu et al., 2024)	Model Sophistication	ML's efficacy is circumscribed by algorithmic complexity and the intricacy of feature sets employed.	DL adeptly tackles high-dimensional data and discerns intricate patterns, owing to its neural architecture.	Such prowess expedites the creation of intricate language models, thus enhancing generative AI capabilities.
(Kukreja et al., 2024; Hadi et al., 2023; Kaur et al., 2024; Pahune & Chandrasekharan, 2023; Xu et al., 2024; Zhang et al., 2023)	Learning Paradigm	ML predominantly relies on supervised learning paradigms bolstered by labeled datasets.	In contrast, DL seamlessly integrates supervised and unsupervised learning modalities, enhancing adaptability.	This augmentation markedly amplifies the assimilation capacity from expansive and diverse datasets, enriching model training.
(Yue et al., 2023; Lappin, 2024; Kalyan, 2023; Alqahtani et al., 2023)	Architectural Advancement	Traditional ML methodologies encompass decision trees, SVMs, and linear regression models.	DL embraces intricate architectures such as CNNs, RNNs, LSTMs, and Transformers.	This underpins the deployment of sophisticated architectures like Transformers, pivotal in models like ChatGPT.
(Hadi et al., 2023; Kukreja et al.,	Scalability	ML's scalability is often impeded	In stark contrast, DL's scalability	This scalability empowers the

2024; Hadi et al., 2023; Jovanovic & Voss, 2024; Xu et al., 2024)		by the limitations associated with large datasets and model sizes.	is significantly heightened through GPU and TPU utilization.	training of colossal models like ChatGPT, exponentially augmenting model capabilities.
(Hadi et al., 2023; Kukreja et al., 2024; Xu et al., 2024; Karanikolas et al., 2023; Alwahedi et al., 2024)	Performance Optimization	ML optimization predominantly hinges on traditional techniques like gradient descent.	DL employs advanced optimization algorithms such as Adam and RMSProp.	Such optimization methodologies catalyze training efficiency and augment model performance manifold.
(Chang et al., 2024; Hadi et al., 2023; Kukreja et al., 2024; Hadi et al., 2023; Kaur et al., 2024; Jovanovic & Voss, 2024)	Versatility and Adaptability	ML exhibits relative inflexibility in accommodating diverse data types and task paradigms.	DL boasts exceptional flexibility, adept at handling varied data types and tasks seamlessly.	This versatility facilitates cross-domain applications, enhancing the utility and adaptability of generative AI models.
(Bharathi Mohan et al., 2024; Yan et al., 2024; Myers et al., 2024; Pahune & Chandrasekharan, 2023)	Innovation Propagation	ML serves as the linchpin for pioneering advanced algorithms and computational frameworks.	DL spearheads innovation in neural architectures and learning paradigms.	This innovation fosters the development of novel generative capabilities and expands the horizons of AI applications, exemplified in models like ChatGPT.
(Hadi et al., 2023; Hadi et al., 2023; Pahune & Chandrasekharan, 2023; Xu et al., 2024; Karanikolas et al., 2023)	Application Domain	ML finds application in predictive analytics, classification, and clustering	DL reigns supreme in domains like image recognition, NLP, and game playing.	These advancements drive natural language understanding and generation capabilities in



domains  
predominantly.

generative  
models such as  
ChatGPT,  
revolutionizing  
AI applications.

---

Reinforcement learning has also played a major role in the progress of generative AI. In reinforcement learning, models improve decision-making by receiving rewards or penalties as feedback for their actions. This cycle of feedback assists the model in enhancing its performance gradually. An example is the utilization of reinforcement learning from human feedback (RLHF) to adjust models such as ChatGPT. During this procedure, the model's answers are evaluated by people, and this evaluation is utilized to refine the model's settings, resulting in interactions that are more precise and similar to human behavior. The continuous cycle of receiving feedback has been crucial in improving the accuracy and dependability of generative AI results. The significance of ethics and responsible AI is growing in the creation and utilization of generative AI. Machine learning is essential in addressing these issues by allowing for the application of ethical standards and methods to reduce bias. Sophisticated machine learning methods are applied to identify and minimize biases in training data, ensuring that generative models generate fair and impartial results. Furthermore, it is crucial for establishing user confidence that AI models are easily understandable and transparent. Explanatory techniques like XAI assist in clarifying the decision-making procedures of intricate models such as ChatGPT, increasing their comprehensibility and reliability for users.

The uses of artificial intelligence that generates content through machine learning are extensive and diverse, covering numerous industries and scenarios. For instance, AI-powered chatbots are changing the way businesses engage with customers in customer support. These chatbots are able to address a variety of questions, giving immediate answers and enhancing customer happiness. Generative AI is utilized in education to develop customized learning experiences, with AI tutors adjusting to students' individual needs to improve their learning results. Moreover, AI models are aiding writers and marketers in content creation by producing creative material, which results in saving time and resources while upholding quality. Machine learning progress has also made it easier to incorporate multiple modes in generative AI models. Multimodal AI includes the capacity to handle text, images, and various forms of data at the same time. This integration improves the capabilities of generative models, enabling them to tackle more complicated tasks involving the comprehension and creation of various types of data. For example, artificial intelligence models are able to produce descriptive text from images or generate visual content from written descriptions. This ability to use multiple modes

allows for new opportunities in areas like healthcare, where AI can help identify medical conditions through the analysis of medical images and the creation of in-depth reports.

Although there has been considerable progress, there are still obstacles to overcome in the development of generative AI. A major worry is the possible abuse of content produced by artificial intelligence, like making deepfakes or spreading misinformation. Continuing research and strong safeguards are needed to address these challenges and prevent harmful uses of AI technology. Furthermore, it is crucial to prioritize the protection and confidentiality of user data, as generative AI models frequently depend on large quantities of personal data for their training and functioning. Enforcing strict data protection protocols and following ethical guidelines are crucial for upholding public confidence in AI technologies. The potential of generative AI in the future, enabled by machine learning, is very promising. With ongoing research pushing limits, we should anticipate AI models becoming even more advanced and capable. The advancement of generative AI could be further sped up by technologies like quantum computing and neuromorphic engineering, allowing it to tackle harder problems and tasks it cannot currently handle. Additionally, the continued collaboration among researchers, industry experts, and policymakers will play a critical role in influencing the future of AI, guaranteeing that its advantages are achieved while reducing possible dangers.

## **Role of deep learning in advancing generative artificial intelligence such as ChatGPT**

### **Deep Learning Foundations**

Deep learning, which is a component of machine learning, depends on deep neural networks - neural networks with multiple layers - to represent intricate patterns in data (Raiaan et al., 2024; Hadi et al., 2023; Alqahtani et al., 2023; Cui et al., 2024). The reason for deep learning's triumph in generative AI lies in its capacity to analyze and produce human-like text using extensive data (Lam et al., 2024; Thirunavukarasu et al., 2023; Wang et al., 2024; Liu et al., 2024; 40). The structure of deep learning models, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformers, offers the necessary computational framework for tasks like NLP (Chang et al., 2024; Liu et al., 2024; Bai et al., 2024; Xu et al., 2024; Zhang et al., 2023). Fig. 5.2 illustrates the process of how deep learning contributes to advancing generative artificial intelligence, such as ChatGPT.

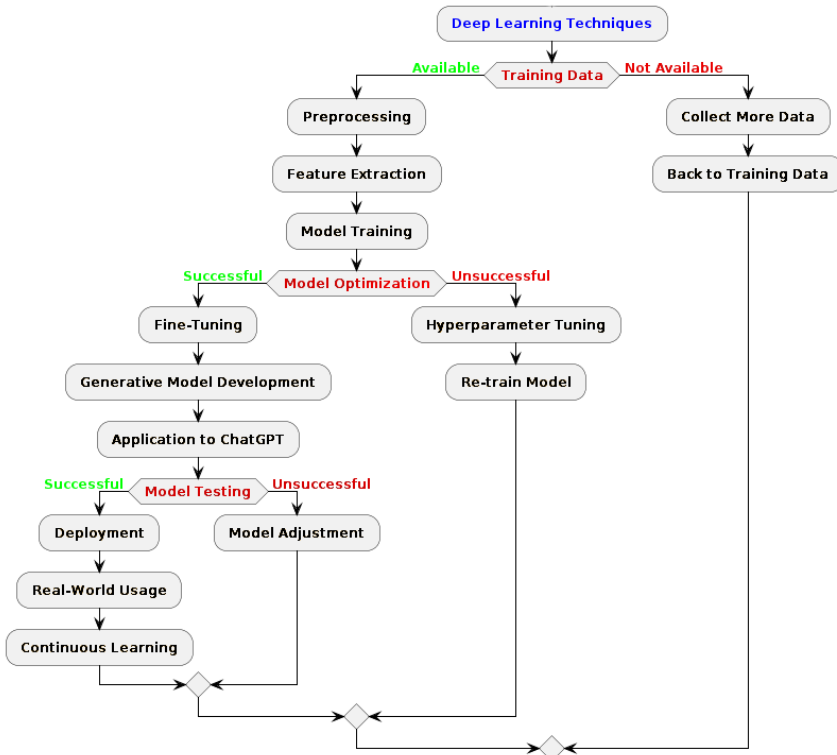
### **Evolution of Language Models**

The evolution of language models started with basic designs such as RNNs, which were able to understand sequential patterns in information. Nevertheless, these models

encountered constraints in managing long-term dependencies because of issues with vanishing gradients. The development of Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs) offered some improvement, but the major advancement occurred with the transformer design. Transformers leverage self-attention mechanisms to determine the significance of various words in a sentence, enabling enhanced comprehension of context and parallel processing in training.

### The Transformer Architecture

The transformer architecture transformed the field of NLP by overcoming the constraints of RNNs and LSTMs. The self-attention mechanism allows the model to concentrate on important sections of the input sequence, regardless of where they are located, leading to great performance in language modeling assignments. Transformers are made up of stacks of encoders and decoders, with the encoder handling the input sequence and the decoder producing the output sequence. This design set the groundwork for future models, such as the Generative Pre-trained Transformer (GPT) series.



**Fig. 5.2** Process of how deep learning contributes to advancing generative artificial intelligence, such as ChatGPT

## **Training Paradigms**

Training generative models such as GPT consists of two primary stages: pre-training and fine-tuning. In pre-training, the model learns how to anticipate the subsequent word in a sentence based on the words that come before it. This stage makes use of unsupervised learning, making use of extensive quantities of unlabelled textual data. Fine-tuning is a type of supervised learning in which the model is trained on particular tasks or datasets to enhance its performance in specific applications. This blend of unsupervised and supervised learning allows the model to produce text that is both grammatically accurate and contextually suitable.

## **Reinforcement Learning and Human Feedback**

An important progress in improving generative models such as ChatGPT involves incorporating reinforcement learning from human feedback (RLHF). RLHF includes educating the model using evaluations from human assessors, who assess the results according to their excellence. This repetitive process assists in aligning the model's results with human preferences, minimizing occurrences of biased or inappropriate reactions. By adding Reinforcement Learning with Human Feedback (RLHF), models such as ChatGPT can generate text that is more dependable and tailored to the user's needs, improving their real-world applicability.

## **Scalability and Computational Resources**

The ability of deep learning models to scale is also a crucial element in the advancement of generative AI. As the need for computational power has increased, along with the size of models and datasets. State-of-the-art hardware, like GPUs and TPUs, is essential for effectively training these large-scale models. Distributed training methods, involving dividing the training workload among several machines, have played a crucial role in handling the computational requirements. Additionally, advancements in software frameworks like TensorFlow and PyTorch have made it easier to create and implement deep learning models. These frameworks offer resources and libraries that simplify the process of designing, training, and optimizing neural networks, enabling researchers and engineers to easily explore and improve generative models.

## **5.4 Conclusions**

The impact of ML and DL on enhancing generative artificial intelligence (AI) models such as ChatGPT is significant and revolutionary. These advancements have led to major progress in creating and improving generative models, leading to interactions that are more human-like and increased abilities. Machine learning, through its capability to

recognize patterns from extensive data sets, forms the fundamental algorithms that support generative artificial intelligence. Deep learning, using neural networks, enables these models to learn and produce intricate language patterns, replicating human dialogue with impressive precision.

Recent developments in transformer designs, such as the GPT series, emphasize the crucial importance of deep learning. These models use huge amounts of data and computing power to learn from various language patterns, leading to a high level of coherence and contextual understanding in text generated by AI. Advancements in training methods like reinforcement learning from human feedback (RLHF) have enhanced the effectiveness of generative AI, leading to better and more valuable results for users. The combining of ML and DL has also made it easier to scale and adapt generative AI. These models can now complete a variety of tasks such as translating languages, creating content, making personalized recommendations, and automating customer service. In addition, current studies on ethical AI and bias reduction are tackling important obstacles to guarantee that generative AI systems are both powerful and acceptable. As research progresses, these technologies are expected to enhance the capabilities of generative AI even more, creating new possibilities in the field of artificial intelligence.

## References

- Alqahtani, T., Badreldin, H. A., Alrashed, M., Alshaya, A. I., Alghamdi, S. S., bin Saleh, K., ... & Albekairy, A. M. (2023). The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research. *Research in Social and Administrative Pharmacy*.
- Alwahedi, F., Aldhaheeri, A., Ferrag, M. A., Battah, A., & Tihanyi, N. (2024). Machine learning techniques for IoT security: Current research and future vision with generative AI and large language models. *Internet of Things and Cyber-Physical Systems*.
- Bai, G., Chai, Z., Ling, C., Wang, S., Lu, J., Zhang, N., ... & Zhao, L. (2024). Beyond efficiency: A systematic survey of resource-efficient large language models. *arXiv preprint arXiv:2401.00625*.
- Bharathi Mohan, G., Prasanna Kumar, R., Vishal Krishh, P., Keerthinathan, A., Lavanya, G., Meghana, M. K. U., ... & Doss, S. (2024). An analysis of large language models: their impact and potential applications. *Knowledge and Information Systems*, 1-24.
- Bridges, L. M., McElroy, K., & Welhouse, Z. (2024). Generative artificial intelligence: 8 critical questions for libraries. *Journal of Library Administration*, 64(1), 66-79.
- Cantens, T. (2024). How will the state think with ChatGPT? The challenges of generative artificial intelligence for public administrations. *AI & SOCIETY*, 1-12.
- Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., ... & Xie, X. (2024). A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology*, 15(3), 1-45.

- Cheng, J. (2024). Applications of Large Language Models in Pathology. *Bioengineering*, 11(4), 342.
- Cui, C., Ma, Y., Cao, X., Ye, W., Zhou, Y., Liang, K., ... & Zheng, C. (2024). A survey on multimodal large language models for autonomous driving. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 958-979).
- Hadi, M. U., Qureshi, R., Shah, A., Irfan, M., Zafar, A., Shaikh, M. B., ... & Mirjalili, S. (2023). A survey on large language models: Applications, challenges, limitations, and practical usage. *Authorea Preprints*.
- Hadi, M. U., Qureshi, R., Shah, A., Irfan, M., Zafar, A., Shaikh, M. B., ... & Mirjalili, S. (2023). Large language models: a comprehensive survey of its applications, challenges, limitations, and future prospects. *Authorea Preprints*.
- Hajikhani, A., & Cole, C. (2024). A critical review of large language models: Sensitivity, bias, and the path toward specialized ai. *Quantitative Science Studies*, 1-22.
- Hsu, Y. C., & Ching, Y. H. (2023). Generative artificial intelligence in education, part one: The dynamic frontier. *TechTrends*, 67(4), 603-607.
- Jovanovic, M., & Voss, P. (2024). Trends and challenges of real-time learning in large language models: A critical review. *arXiv preprint arXiv:2404.18311*.
- Kalota, F. (2024). A Primer on Generative Artificial Intelligence. *Education Sciences*, 14(2), 172.
- Kalyan, K. S. (2023). A survey of GPT-3 family large language models including ChatGPT and GPT-4. *Natural Language Processing Journal*, 100048.
- Karanikolas, N., Manga, E., Samaridi, N., Tousidou, E., & Vassilakopoulos, M. (2023). Large Language Models versus Natural Language Understanding and Generation. In *Proceedings of the 27th Pan-Hellenic Conference on Progress in Computing and Informatics* (pp. 278-290).
- Kaur, P., Kashyap, G. S., Kumar, A., Nafis, M. T., Kumar, S., & Shokeen, V. (2024). From Text to Transformation: A Comprehensive Review of Large Language Models' Versatility. *arXiv preprint arXiv:2402.16142*.
- Kukreja, S., Kumar, T., Purohit, A., Dasgupta, A., & Guha, D. (2024). A literature survey on open source large language models. In *Proceedings of the 2024 7th International Conference on Computers in Management and Business* (pp. 133-143).
- Lam, H. Y. I., Ong, X. E., & Mutwil, M. (2024). Large language models in plant biology. *Trends in Plant Science*.
- Lappin, S. (2024). Assessing the strengths and weaknesses of Large Language Models. *Journal of Logic, Language and Information*, 33(1), 9-20.
- Liu, J., Yang, M., Yu, Y., Xu, H., Li, K., & Zhou, X. (2024). Large language models in bioinformatics: applications and perspectives. *arXiv preprint arXiv:2401.04155*.
- Liu, X., Xu, P., Wu, J., Yuan, J., Yang, Y., Zhou, Y., ... & Huang, F. (2024). Large language models and causal inference in collaboration: A comprehensive survey. *arXiv preprint arXiv:2403.09606*.
- Liu, Y., Cao, J., Liu, C., Ding, K., & Jin, L. (2024). Datasets for Large Language Models: A Comprehensive Survey. *arXiv preprint arXiv:2402.18041*.

- Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., ... & Ge, B. (2023). Summary of chatgpt-related research and perspective towards the future of large language models. *Meta-Radiology*, 100017.
- Min, B., Ross, H., Sulem, E., Veyseh, A. P. B., Nguyen, T. H., Sainz, O., ... & Roth, D. (2023). Recent advances in natural language processing via large pre-trained language models: A survey. *ACM Computing Surveys*, 56(2), 1-40.
- Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., & Gao, J. (2024). Large language models: A survey. *arXiv preprint arXiv:2402.06196*.
- Myers, D., Mohawesh, R., Chellaboina, V. I., Sathvik, A. L., Venkatesh, P., Ho, Y. H., ... & Jararweh, Y. (2024). Foundation and large language models: fundamentals, challenges, opportunities, and social impacts. *Cluster Computing*, 27(1), 1-26.
- Omiye, J. A., Gui, H., Rezaei, S. J., Zou, J., & Daneshjou, R. (2024). Large Language Models in Medicine: The Potentials and Pitfalls: A Narrative Review. *Annals of Internal Medicine*, 177(2), 210-220.
- Pahune, S., & Chandrasekharan, M. (2023). Several categories of large language models (llms): A short survey. *arXiv preprint arXiv:2307.10188*.
- Park, Y. J., Pillai, A., Deng, J., Guo, E., Gupta, M., Paget, M., & Naugler, C. (2024). Assessing the research landscape and clinical utility of large language models: A scoping review. *BMC Medical Informatics and Decision Making*, 24(1), 72.
- Raiaan, M. A. K., Mukta, M. S. H., Fatema, K., Fahad, N. M., Sakib, S., Mim, M. M. J., ... & Azam, S. (2024). A review on large Language Models: Architectures, applications, taxonomies, open issues and challenges. *IEEE Access*.
- Shi, H., Xu, Z., Wang, H., Qin, W., Wang, W., Wang, Y., & Wang, H. (2024). Continual Learning of Large Language Models: A Comprehensive Survey. *arXiv preprint arXiv:2404.16789*.
- Su, J., Jiang, C., Jin, X., Qiao, Y., Xiao, T., Ma, H., ... & Lin, J. (2024). Large Language Models for Forecasting and Anomaly Detection: A Systematic Literature Review. *arXiv preprint arXiv:2402.10350*.
- Thirunavukarasu, A. J., Ting, D. S. J., Elangovan, K., Gutierrez, L., Tan, T. F., & Ting, D. S. W. (2023). Large language models in medicine. *Nature medicine*, 29(8), 1930-1940.
- Tustumi, F., Andreollo, N. A., & Aguilar-Nascimento, J. E. D. (2023). Future of the language models in healthcare: the role of chatGPT. *ABCD. arquivos brasileiros de cirurgia digestiva (são paulo)*, 36, e1727.
- Veres, C. (2022). Large language models are not models of natural language: they are corpus models. *IEEE Access*, 10, 61970-61979.
- Wang, J., Wu, Z., Li, Y., Jiang, H., Shu, P., Shi, E., ... & Zhang, S. (2024). Large language models for robotics: Opportunities, challenges, and perspectives. *arXiv preprint arXiv:2401.04334*.
- Wang, S., Xu, T., Li, H., Zhang, C., Liang, J., Tang, J., ... & Wen, Q. (2024). Large language models for education: A survey and outlook. *arXiv preprint arXiv:2403.18105*.
- Wei, C., Wang, Y. C., Wang, B., & Kuo, C. C. J. (2023). An overview on language models: Recent developments and outlook. *arXiv preprint arXiv:2303.05759*.
- Xie, Q., Schenck, E. J., Yang, H. S., Chen, Y., Peng, Y., & Wang, F. (2023). Faithful ai in medicine: A systematic review with large language models and beyond. *medRxiv*.

- Xu, H., Wang, S., Li, N., Zhao, Y., Chen, K., Wang, K., ... & Wang, H. (2024). Large language models for cyber security: A systematic literature review. arXiv preprint arXiv:2405.04760.
- Xu, X., Xu, Z., Ling, Z., Jin, Z., & Du, S. (2024). Emerging Synergies Between Large Language Models and Machine Learning in Ecommerce Recommendations. arXiv preprint arXiv:2403.02760.
- Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., ... & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, 55(1), 90-112.
- Yang, R., Tan, T. F., Lu, W., Thirunavukarasu, A. J., Ting, D. S. W., & Liu, N. (2023). Large language models in health care: Development, applications, and challenges. *Health Care Science*, 2(4), 255-263.
- Yue, T., Wang, Y., Zhang, L., Gu, C., Xue, H., Wang, W., ... & Dun, Y. (2023). Deep learning for genomics: from early neural nets to modern large language models. *International Journal of Molecular Sciences*, 24(21), 15858.
- Zhang, Z., Fang, M., Chen, L., Namazi-Rad, M. R., & Wang, J. (2023). How do large language models capture the ever-changing world knowledge? a review of recent advances. arXiv preprint arXiv:2310.07343.
- Zhao, H., Chen, H., Yang, F., Liu, N., Deng, H., Cai, H., ... & Du, M. (2024). Explainability for large language models: A survey. *ACM Transactions on Intelligent Systems and Technology*, 15(2), 1-38.
- Zhuhadar, L. P., & Lytras, M. D. (2023). The application of AutoML techniques in diabetes diagnosis: current approaches, performance, and future directions. *Sustainability*, 15(18), 13484.