A Brief Introduction to Generative AI and Business Enhancement Solutions

Authors: Gu Liping, Gu Liping

Date: 2023-10-18T00:00:00+00:00

Abstract

Artificial intelligence encompasses machine learning, deep learning, generative artificial intelligence, natural language processing, computer vision, reinforcement learning, among others. Generative artificial intelligence represents a branch of artificial intelligence capable of generating novel data, with foundational theories including sparse coding, autoencoders, and brain-inspired generative artificial intelligence. At its core are generative models that achieve the objective of producing new data through deep learning methodologies such as generative adversarial networks (GANs), variational autoencoders (VAEs), and generative pre-trained transformer models (e.g., GPT). Generative artificial intelligence (GAI) finds extensive applications in natural language processing, computer vision, audio processing, creative domains, and various other fields. GAI models exhibit broad application potential across diverse domains, driving innovation and exerting profound influence on industries worldwide.

Full Text

Preamble

A Brief Introduction to Generative Artificial Intelligence and Business Improvement Solutions

Gu Liping[1][2]

[1] National Science Library, Chinese Academy of Sciences, Beijing, 100190
[2] Department of Information Resources Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing, 100190

Abstract
Artificial intelligence encompasses machine learning, deep learning, generative artificial intelligence, natural language processing, computer vision, reinforcement learning, and other subfields. Generative artificial intelligence represents

a branch of AI capable of producing novel data. Its foundational theories include sparse coding, autoencoders, and brain-inspired generative AI. The core of generative AI lies in generative models, which achieve the objective of creating new data through deep learning methods such as generative adversarial networks (GANs), variational autoencoders (VAEs), and generative pre-trained transformer models (e.g., GPT). Generative artificial intelligence (GAI) finds extensive applications across natural language processing, computer vision, audio processing, creative domains, and numerous other fields. GAI models exhibit broad application potential across various domains, driving innovation and exerting profound influence on diverse industries.

Keywords: Generative Artificial Intelligence; AI-Generated Content (AIGC); AI for Science; AI for Library
Classification Number: G250

1. Development Domains of Artificial Intelligence

Artificial intelligence (AI) constitutes a broad field encompassing numerous subdomains and technologies, which have given rise to various interdisciplinary application branches across academic disciplines.

1.1 Machine Learning

Machine learning represents a subdomain of AI that employs algorithms enabling computer systems to learn from and improve through data. This includes supervised learning, unsupervised learning, reinforcement learning, and various other methodologies. As a core technology of AI, machine learning empowers computers to automatically learn and improve from data without explicit programming.

1.2 Deep Learning

Deep learning, a branch of machine learning, utilizes neural network models to process and comprehend complex data such as images, speech, and natural language. As a crucial subfield of AI, deep learning has achieved remarkable success in numerous domains due to its capacity to handle intricate data patterns.

1.3 Natural Language Processing (NLP)

Natural language processing focuses on enabling computers to understand and generate human language, encompassing tasks such as text processing, text generation, and machine translation. This represents a vital AI branch because it facilitates human-computer linguistic interaction.

1.4 Computer Vision

Computer vision aims to enable computers to understand and interpret images and video, involving tasks such as object recognition, image classification, and

target detection. This subdomain is essential as it grants machines visual perception capabilities.

1.5 Reinforcement Learning

Reinforcement learning is a machine learning paradigm where agents learn to make decisions through interaction with an environment to maximize cumulative rewards. This approach is significant because it enables computers to learn optimal behaviors via environmental feedback.

These domains constitute integral components of artificial intelligence that interconnect to form a complete AI ecosystem. For instance, deep learning and reinforcement learning are both subdomains of machine learning, while natural language processing and computer vision represent AI applications in specific domains.

2. Foundational Theories of Generative Artificial Intelligence

Generative Artificial Intelligence (GAI) refers to AI technologies capable of generating new data, images, text, or other types of information. As a branch of artificial intelligence, GAI leverages AI techniques to create novel data and content.

The foundational theories of generative AI include probabilistic graphical models, autoencoders, and generative adversarial networks. Probabilistic graphical models represent dependencies among random variables using probability-based structures. Autoencoders are neural network models that learn low-dimensional representations of data. Generative adversarial networks consist of a generator and discriminator that collaboratively produce new data. These theoretical frameworks enable GAI models to generate novel data samples rather than merely classifying or regressing existing data.

Probabilistic graphical models, which include directed models (e.g., Bayesian networks, Markov models) and undirected models (e.g., Markov random fields), can represent various data types including images, text, and speech. Autoencoders comprise an encoder that maps input data to a low-dimensional latent representation and a decoder that reconstructs the original data from this representation, making them suitable for data compression, denoising, and feature extraction. Generative adversarial networks employ a generator that attempts to produce realistic data and a discriminator that distinguishes between real and generated data, enabling the creation of images, audio, and text.

These theoretical foundations empower generative AI to create new samples across modalities. For example, generative adversarial networks can synthesize novel images and audio, while autoencoders can generate new text data.

3. Core Technologies of Generative Artificial Intelligence

The engineering implementation of generative AI involves multiple core technologies, including deep learning, generative models, natural language processing, computer vision, and reinforcement learning, which collectively provide the foundation for generative tasks.

3.1 Deep Learning

Deep learning simulates the neural structure of the human brain through multilayered neuron networks that extract features and model complex relationships from data. As the foundational technology for generative AI, deep learning provides powerful tools for building generative models and implementing various generative tasks.

3.2 Generative Models

Generative models simulate data distributions to produce novel data samples, including generative adversarial networks (GANs), variational autoencoders (VAEs), and autoregressive models. These constitute the core of generative AI, enabling the creation of diverse data types such as images, text, and audio.

Generative adversarial networks (GANs), variational autoencoders (VAEs), and autoregressive models represent commonly used generative models in GAI. Their fundamental principles include:

(1) Generative Adversarial Networks (GANs): GANs comprise two neural networks—a generator and a discriminator—that work together through adversarial training. The generator attempts to produce samples resembling real data, while the discriminator tries to distinguish generated samples from authentic data.

Generator: $G(z)$, where z is a random noise vector sampled from the latent space.
Discriminator: $D(x)$, where x is an input data sample.

Generator loss function: $L_G = \mathbb{E}[\log(1 - D(G(z)))]$.
Discriminator loss function: $L_D = -[\mathbb{E}[\log(D(x))] + \mathbb{E}[\log(1 - D(G(z)))]]$.

(2) Variational Autoencoders (VAEs): VAEs integrate the encoder-decoder architecture of autoencoders while introducing latent variables z to learn data probability distributions. VAEs aim to learn latent representations and generate new samples.

Encoder: $Q(z|x)$ represents the posterior distribution of latent variable z given input x.
Decoder: $P(x|z)$ represents the conditional distribution for generating data x given latent variable z.

Prior distribution of latent variable z: $P(z)$.
Objective function (variational lower bound): $L = \mathbb{E}[\log P(x|z)] - \mathrm{KL}(Q(z|x)\|P(z))$.

(3) Autoregressive Models: These models typically represent data as a chained product of conditional probability distributions:

$$P(x) = \prod_t P(x_t|x_1, x_2, ..., x_{t-1})$$

where x denotes a data sequence and x_t represents the t-th element in the sequence.

These generative models find widespread application in GAI for producing different data types. GANs generate high-quality samples through adversarial training, VAEs learn latent representations via probabilistic distributions, and autoregressive models generate sequential data. These models play crucial roles in numerous applications.

3.3 Natural Language Processing (NLP)

NLP technologies enable the processing and generation of natural language text, facilitating applications such as text generation, machine translation, text summarization, and sentiment analysis. These capabilities are closely integrated with generative AI for dialogue systems and text generation tasks.

3.4 Computer Vision

This domain involves processing and generating image and video data, including image generation, super-resolution, and segmentation. Techniques such as GANs for image generation and style transfer exemplify computer vision's role in generative AI.

3.5 Reinforcement Learning

Reinforcement learning enables agents to learn optimal actions within an environment to maximize cumulative rewards, which can be applied to optimize generation strategies for tasks such as image captioning and automated machine translation.

These core technologies are essential for implementing generative AI. Generative models and deep learning methods model data distributions and produce novel samples, while NLP and computer vision technologies handle text and image generation tasks. Reinforcement learning can optimize generation strategies to achieve superior results. Collectively, generative AI represents an application domain that leverages these technologies to create new content and data, demonstrating extensive potential across industries.

4. Application Domains of Generative Artificial Intelligence

Generative artificial intelligence finds application across numerous domains, including but not limited to:

4.1 Text Generation

GAI can generate novel textual data, such as automatic news reporting, advertising copywriting, and virtual character dialogue.

4.2 Image Generation

GAI can synthesize new image data, including human faces, landscapes, and other visual content.

4.3 Audio Generation

GAI can create new audio data, such as music composition and speech synthesis.

4.4 Data Privacy Protection

GAI can generate synthetic data to protect sensitive information, such as synthetic medical records or financial data.

4.5 Creative Industries

Generative AI is widely applied in arts, design, and creative sectors for producing artworks, music, literature, designs, and virtual characters.

4.6 Data Augmentation

GAI enhances training datasets to improve performance on supervised learning tasks such as image classification and text categorization.

4.7 Simulation and Modeling

Generative models simulate natural phenomena, including weather forecasting, traffic simulation, and virtual environments.

These application domains represent only a subset of GAI's potential. As technology continues to evolve, the scope of generative AI applications will expand further.

4.8 Library Applications

Generative AI can enhance efficiency in library operations, improve user experience, and optimize information resource management. Extensive research has explored this potential, with key applications including:

Document Summarization: Utilizing generative models such as GPT-4 or iFLYTEK to automatically generate summaries of documents or books, enabling users to quickly grasp core content and save time.

Intelligent Search: Developing intelligent search engines using natural language processing techniques to provide more accurate results based on user queries, thereby improving the search experience.

Automatic Classification and Tagging: Employing computer vision technology to automatically classify and tag books, documents, and other materials, facilitating organized collections and improved resource accessibility.

Speech Recognition and Transcription: Converting spoken documents and notes into text using speech recognition technology, enhancing document searchability and eliminating manual transcription effort.

Personalized Recommendation: Leveraging machine learning and collaborative filtering algorithms to provide personalized book and document recommendations, helping users discover relevant resources more efficiently.

Automatic Index and Table of Contents Generation: Using generative models to automatically create indexes and tables of contents for books or documents, reducing catalogers' workload while ensuring accuracy.

Automatic Reference Generation: Employing generative models to automatically produce academic references, assisting scholars and authors in creating accurate bibliographies.

Intelligent Question-Answering Systems: Implementing systems that automatically generate answers or provide relevant document links in response to user questions, improving information retrieval efficiency.

Data Cleaning and Archive Management: Using natural language processing techniques to automatically detect and correct errors and redundancies in documents before archiving, thereby maintaining high data quality.

By applying generative AI technologies to library operations, institutions can enhance resource management, information retrieval, and user services while providing improved user experiences and resource availability.

References

[1] Alto V. *Modern Generative AI with ChatGPT and OpenAI Models*. Birmingham: Packt Publishing, 2023.
[2] Rothman D. *Transformers for Natural Language Processing: Build, train, and fine-tune deep neural network architectures for NLP with Python, PyTorch, TensorFlow, BERT, and GPT-3*. Birmingham: Packt Publishing, 2022.
[3] Che Wanxiang, Guo Jiang. *Natural Language Processing: Methods Based on Pre-trained Models* (in Chinese). Beijing: Publishing House of Electronics Industry, 2021.

[4] Saito Koyuki, Lu Yujie (trans.). *Deep Learning Advanced: Natural Language Processing* (in Chinese). Beijing: Posts & Telecom Press, 2020.

Note: Figure translations are in progress. See original paper for figures.