The Pervasive Influence of Artificial Intelligence in the IT Industry

Bhuvaneswari S
Department of Computer Science and Engineering
Easwari engineering college,
Chennai, India
bhuvaneswari.s@eec.srmrmp.edu.in

Deepakraj R
Department of Computer Science and Engineering
Easwari engineering college
Chennai, India
deepakrajmohan3602@gmail.com

ABSTRACT

Artificial Intelligence is more than just a buzzword; it is a game-changer that genuinely aids and significantly improves the IT industry. Right here, right now, AI is making a tangible impact, revolutionizing the way we work and interact with technology. So, why not embrace its transformative potential in the IT industry? This paper presents a comprehensive review of the usage of AI in the IT industry, highlighting its diverse applications, benefits, and future implications. The rapid advancement of Artificial Intelligence (AI) has revolutionized the Information Technology (IT) industry, impacting various aspects of business operations and transforming the way technology is leveraged. The study encompasses advanced AI technologies like Federated Learning (FL), Generative AI (GAI), and Explainable AI (XAI), exploring their integration across various IT departments and domains. The paper delves into the significance of enabling AI in IT industry, ethical considerations of AI adoption, emphasizing the need for responsible AI practices. The paper concludes with highlighting the future potential of those advanced AI technologies and their limitations.

Keywords—artificial intelligence; it industry; federated learning; generative ai; explainable ai; machine learning; deep learning; computer vision; natural language processing; automation

I. INTRODUCTION

This is the enticing realm of the information technology, where innovation and development meet to create the digital environment of the contemporary world. As we move deeper into the twenty-first century, the IT industry has developed into an exciting force propelling advancement in every aspect of human existence. The present IT sector pulses with potential that redefines the limits of human performance, from cutting-edge artificial intelligence altering commercial operations to the seamless interconnectedness brought about by the Internet of Things. The IT industry is the ultimate change agent in this dynamic environment, inspiring businesses to embrace digital transformation and individuals to realize their full potential. It also acts as a backbone of our digital age, enabling us to transcend boundaries, communicate with one another, and imagine a future that knows no bounds, from elegant smartphones that put the world at our fingertips to ground-breaking data analytics boosting decision-making. The contemporary IT industry promises a world of marvels still to be discovered as we set out on this fascinating adventure, where innovation and ingenuity come together to shape the landscape of the future. The core of the IT sector lies in a dynamic ecosystem that is constantly powered by brilliant brains and game-changing innovations. The convergence of many disciplines, such as Artificial Intelligence (AI), Cloud computing, Cyber Security, and the Internet of Things, has triggered a technological revolution that is transforming the fabric of societies worldwide.

As we go across the arena of today's IT business, we discover the enthralling terrain that has permanently changed our lives. It is regarded as the modern world's backbone, playing a critical role in determining how we live, work, and communicate. It has evolved from a specialized industry to a critical driver of global prosperity and innovation. An IT industry become a driving force behind economic growth, generating significant money and creating millions of employment worldwide [1]. Beyond its commercial influence, the IT sector is viewed as a societal enabler, transforming healthcare, education, transportation, and other key sectors. As the world's reliance on digital solutions grows, the IT sector's importance grows, establishing itself as a crucial actor in building a more connected and technology-driven future.

II. SYSTEMATIC REVIEW OF THE CURRENT STATE OF IT INDUSTRY

The information technology industry is a multidimensional sector that acts as both a commercial entity and a potential force for the progress of society. From a business standpoint, the IT industry is a significant economic contributor, producing significant revenue and job possibilities globally. Today most of the IT organizations, ranging from small startups to international corporations, promote innovation, develop cutting-edge technology, and provide businesses and consumers with a diverse range of products and services [2].

The quantitative contribution of the IT industry to society can be observed by its impact on numerous sectors, for example, it has enabled substantial breakthroughs in healthcare through electronic health records, telemedicine, and medical imaging technologies. According to research published in the Journal of Medical Internet Research, telemedicine saves an average of $19 to $121 every patient visit by eliminating the need for in-person appointments and improving healthcare accessible. IT has altered learning approaches in education, enabling e-learning platforms and tailored learning experiences. According to a US Department of Education survey, 93% of teachers use edtech in their classes, which dramatically improves student engagement and academic outcomes [3].

Furthermore, the IT industry has played a significant role in fostering societal changes through environmental sustainability, for example, Use of cloud computing has considerably reduced energy usage and carbon emissions. According to a study conducted by the Lawrence Berkeley National Laboratory, cloud services could lower energy consumption by up to 87% for specific applications when compared to traditional IT infrastructure. Furthermore, smart grid systems powered by IT have improved energy efficiency and reduced energy waste, resulting in lower greenhouse gas emissions. In addition, IT firms frequently participate in corporate social responsibility (CSR) programs, which assist social issues and community development. According to the Corporate Giving Standard, 69% of IT companies in the United States donated to charitable causes in 2020. These gifts make a real difference, from assisting local communities to sponsoring education initiatives and disaster relief activities.

Besides the advancements and differences made by the IT industry to society, it also significantly contributes to a nation's GDP growth. In the Fiscal year 2022, the IT industry contributed the share of about 9.3% to the GDP of the United States, 7.4% to the GDP of India, 4% to the GDP of China, and 9% to the GDP of the European Union, as per the study [4][5][7][8]. In addition to, the contribution of IT sector to the GDP of US will be expected to reach about 10% in 2025. The IT industry is a dynamic and globalized sector, with countries collaborating and competing to the technological progress and innovation. Many nations such as India, United States, China, Japan, Taiwan, Germany contribute significantly to technological advancements, each with its unique strengths and areas of expertise. Among all India has emerged as a prominent player in the global IT industry, playing a significant role in various aspects of the sector. The country's IT prowess has been driven by its skilled workforce, a vast pool of technical talent, cost-effectiveness, and a conducive business environment [6]. It gave a consistent result on contributing to the country’s GDP growth in past five years ranging from 7.4 to 8 percent as shown in Figure 1.

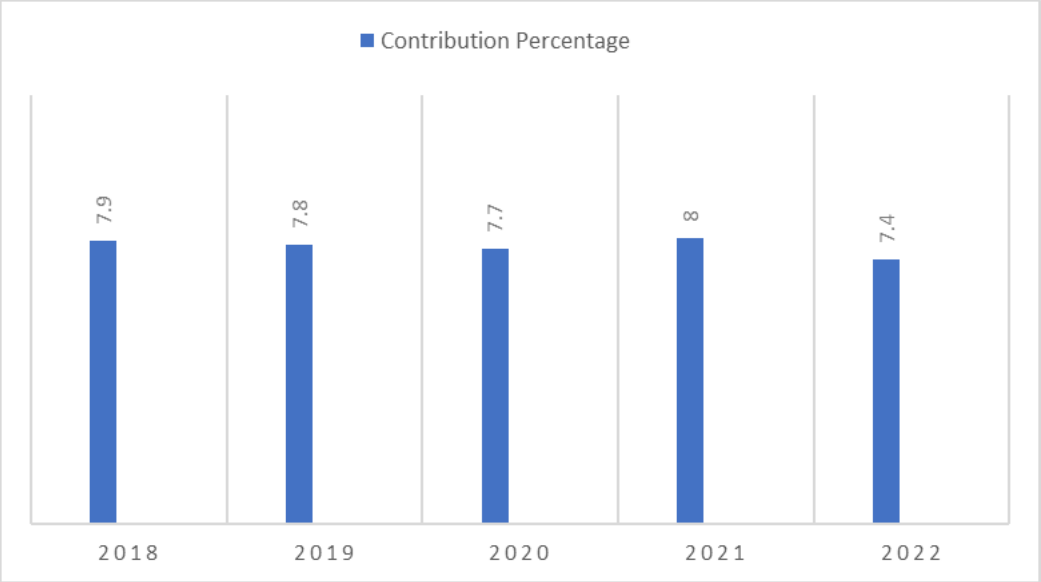

Figure 1: Consistent share of IT industry to the growth of India's GDP in past five years from 2018 to 2022

A. **Domination of AI**

The IT sector relies on a wide range of technologies such as cloud computing, artificial intelligence, Internet of Things, Software development frameworks and much more to facilitate various operations, improve efficiency, and drive innovation. In the above list, while other technologies are vital in their respective domains, AI's versatility and transformative power make it an accelerator in reshaping the IT industry and facilitating the innovation across various sectors. AI has emerged as a dominant force in the IT industry, transforming the way businesses operate and interact with technology. One of the key reasons for AI's domination is its ability to drive automation and efficiency across various IT processes. AI-powered tools and algorithms streamline software development, data analysis, and IT operations, leading to faster, more accurate results and reduced human effort.

The applicability of AI across multiple domains have contributed to its dominance. AI has enabled achievements in domains such as virtual assistants, chatbots, recommendation systems, and autonomous vehicles, ranging from natural language processing (NLP) to computer vision and robotics [9]. The capacity of AI to continuously learn and develop from data ensures its applicability to a wide range of industry-specific use cases, spurring innovation and offering up new business opportunities. As AI technologies progress and become more widely available, their integration into the IT industry is projected to deepen further, redefining business models and revolutionizing how technology is used to meet the changing needs of the modern world. Here are some of the key departments where AI is applied to enhance the progress and improving efficiency:

- Software Development
- Code Integration
- Information Retrieval and Search Engines
- Data Analytics
- Business Intelligence
- Fintech
- Education and E-learning
- Recommendation Systems
- IoT Integration
- Research and Development
- Automated Testing and Validation

B. **Usage of AI Methods in IT**

Being a superset, AI contains lot methods and techniques that can be applied or implemented for various specific purposes. By categorizing AI methodologies, enterprises may find the best effective approach for their specific IT activities and harness AI's potential to improve operations, and achieve better business outcomes. Accordingly, AI technologies are mainly comprising into four different categories, they are Machine learning and Deep learning, Natural Language Processing (NLP), Robotics and Automation, and Computer Vision.

1) Machine Learning and Deep Learning

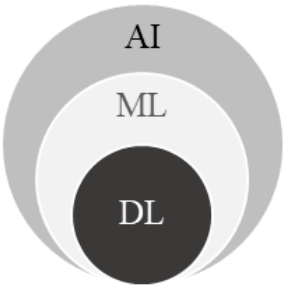

Figure 2: Illustration of AI, ML and DL with the concept of subsets

Deep learning is a subset or core of Machine learning whereas Machine learning is a subset of AI (Figure 2). Both have significant usage in the IT industry, revolutionizing various aspects of technology and business operations. ML techniques, such as supervised and unsupervised learning, are applied in tasks like data analysis, and predictive modelling. DL involves training a computer system to do classification tasks directly from sounds, texts, or images utilizing a vast amount of data [10].

2) Natural Language Processing

Figure 3: Illustration of the concepts in Natural Language Processing (NLP)

Natural Language Processing (NLP) is critical in the IT sector for enabling computers to understand, interpret, and interact with human language. Sentiment analysis, chatbots and virtual assistants for customer assistance, language translation, speech recognition, text summarization, and information extraction are just a few of the applications for NLP (Figure 3). It improves user experiences by allowing users to communicate with software and systems using natural language, and it enables organizations to acquire useful insights from unstructured text data. NLP provides effective data analysis, decision-making, and language-related task automation, ultimately transforming how information is handled, transmitted, and used in the IT industry.

3) Robotics and Automation

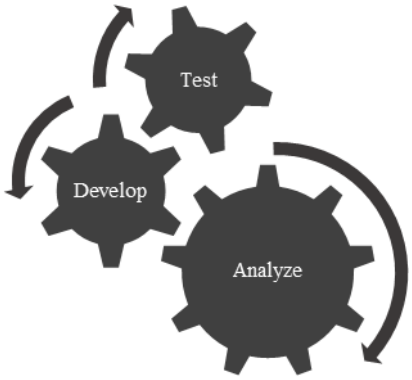

Figure 4: Illustration of automation concept in IT industry

Robotics and automation are critical in the IT business for optimizing operations, increasing efficiency, and eliminating human interference. RPA is commonly used to automate repetitive and rule-based operations like data input, report production, and workflow management, freeing up human resources for more strategic and creative work. In software development, testing, and deployment, automation tools and frameworks are used, resulting in faster and more reliable application delivery (Figure 4). Furthermore, automation is used in IT operations for tasks such as infrastructure provisioning, configuration management, and monitoring to provide seamless and efficient IT administration. Robotics and automation have become essential components in the information technology industry, increasing efficiency and allowing firms to focus on higher-value jobs and innovation.

4) Computer Vision

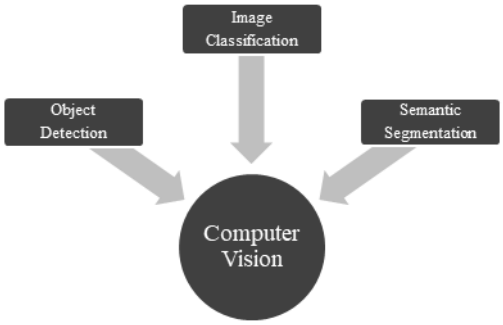

Figure 5: Illustration of the main focus of computer vision

In the IT industry, computer vision is widely utilized to enable machines to analyze and grasp visual information from photographs and videos. Image recognition, object detection, facial recognition, and video analysis are among its many applications (Figure 5). Computer vision is essential in jobs such as automated quality control in manufacturing, surveillance and security systems, self-driving cars, medical image analysis, and augmented reality experiences. The IT industry may extract useful insights from visual data, improve user experiences, and automate numerous processes by employing computer vision technology, resulting in enhanced productivity and innovation across a wide range of domains.

III. THE EVOLVING LANDSCAPE OF AI IN IT

It is critical to embrace future breakthroughs in order to remain competitive in a quickly changing world and to design a brighter future for humanity by leveraging the potential of cutting-edge technologies to address important social, environmental, and economic issues. The future advancements in AI hold immense potential to reshape industries, revolutionize technology, and profoundly impact our lives. Such advancements are expected, where AI systems will possess human-like cognitive capabilities, including reasoning, problem-solving, and creativity. Some of the anticipated advancements of AI are Federated Learning (FL), Generative AI (GAI), and Explainable AI (XAI).

A. **Federated Learning**

FL is a secure machine learning framework that enables ML models to train on decentralized data sources without concession of privacy and security. In traditional machine learning, all data are collected, stored and processed in single location and this can create privacy and security issues. Whereas, the FL enables multiple parties to train the models on their local data and shared the learned parameters to central server. This server aggregates the models followed by updating the global model and preserving the data privacy and data security. Define N Participants in FL, all wants {P1, P2, P3…P_N} to merge the data to train a global model. The Total data $P = P_1 \cup P_2 \ldots \cup P_N$ to train a model T_{sum} with the performance of $P_{sum,}$ where participants co-trained a common model T_{fed} with a performance of P_{fed}. That the no participant exposes its private data to others. Let ε be non-negative, and then the performance loss of FL model is expressed as:

$$| P_{fed} - P_{sum} | < \varepsilon \quad (1)$$

To achieve the learning process of FL by minimizing the loss function, this is calculated by weighted aggregation method

$$WA = w1x1 + w2x2 + w3x3 + w4x. \quad (2)$$

Where w is weighted value, x is input value. The goal of FL is to minimize the objective function,

$$\min o(w) = \sum_{c=1}^{N} \frac{n_c}{n} o_c(w) \quad (3)$$

Where, N represents the number of clients, n_c is the amount of data on the c^{th} client, $o_c(w)$ is the local objective function of the c^{th} client. As described in Figure 6, the following steps constituted the basic framework process of Federated Learning [18],

1. Each client downloads the common model from central server for training.
2. Each client trains the common model with their own local data.
3. The client encrypts their model parameters and uploads in central server for further processing.

Then the central server aggregates the uploaded parameters using FL algorithm and updates the global model.

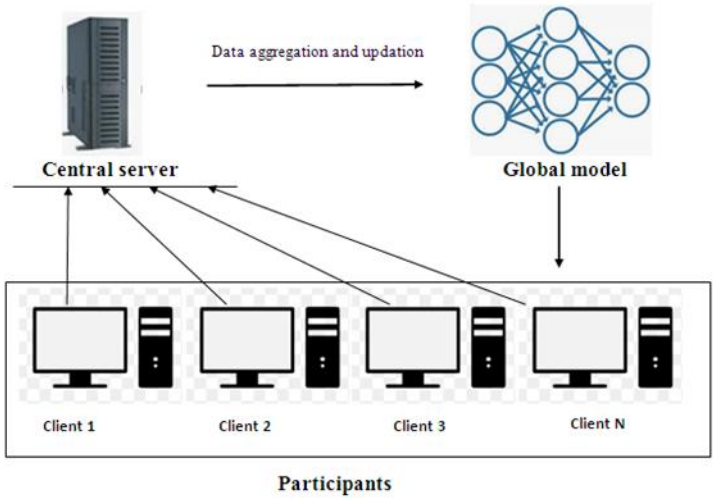

Figure 6: Framework of Federated Learning (FL)

1) The Cataloguing of Federated Learning

The FL is the solution in today's world that offers user data privacy and security. Overtime the FL has taken various formation and shapes in the computing world. The Learning is classified in two ways,

- Types based on schemes
- Types based on Data partitions

a) Types based on schemes

Cross-Silo Federated Learning Model, a silo in information technology is segregated data storage for an association with unstructured raw data. The information is not readily available for further processing due to the restricted access. The silos are connected to a central server to make up FL architecture.

Cross-Device Federated Learning Model, it deals with the data from the individual devices to train the model locally and the central server summative them to create global model.

b) Types based on Data partitions

Horizontal Federated Learning, it also known as Homogenous Federated Learning deals with shared feature space between various participants but sample for each one is different. HFL works on multiple clients within the network. Initially, a local model consisting of typical features is trained locally and communal with the central server, creating a global model.

Vertical Federated Learning Model, it also known as Heterogeneous Federated Learning deals with different features but shares a common sample space as shown in figure 7.

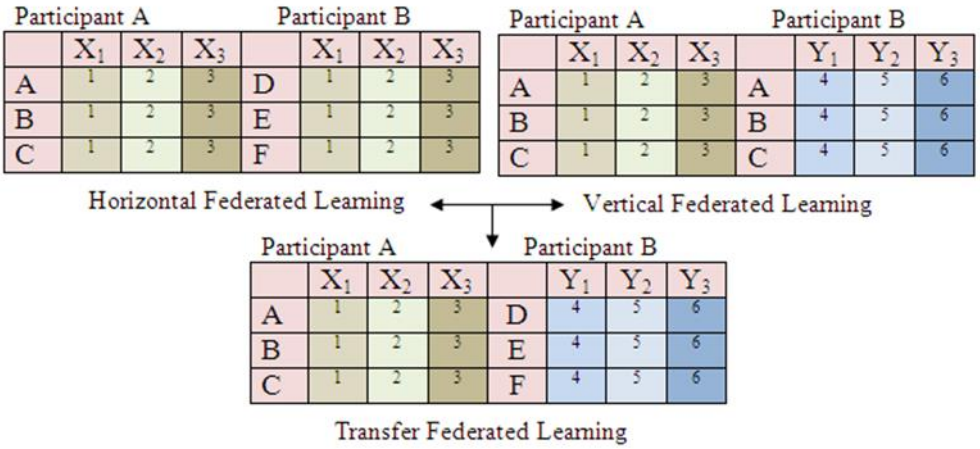

Figure 7: Categorization of Federated Learning (FL)

HFL= Same sample space (X_1, X_2, X_3) but with different features PA= (A, B, C) PB= (D, E, F)
VFL= Different sample space PA= (X_1, X_2, X_3), PB= (Y1, Y_2, Y_3) but with same features (A, B, C)
Federated Transfer Learning Model integrates both HFL and VFL systems with information exchanges across the network without standard features and samples.
TFL= Different sample space PA= (X_1, X_2, X_3) PB= (Y1, Y_2, Y_3) and with different features PA= (A, B, C) PB= (D, E, F)

2) FL Strategies for Privacy Protection

FL is largely guaranteed by security encryption technology with many strategies used in federated learning for protecting the data from leakage [17]. At present, the traditional encryption techniques are securing multi-party computing (SMC), differential privacy (DP), and homomorphic encryption (HE).There are many encryption strategies used in FL will be briefly introduced as follows,

Strat1: SGD - Stochastic Gradient Descent

This technique travels down the gradient of a function on each iteration to converge the gradient with its minimum value. The central server collects all these gradients from multiple clients and averages them. The SGD is the slow processing data and represented as,

$$y = y - \frac{1}{sample}\sum\frac{1}{k}\sum_{k=1}^{k} gradient(x)$$

Strat2: FedAvg-Federated Averaging

This approach is used to implement a cross silo federated learning model which allows the participants/clients to take multiple gradients to an estimated value. Initially, calculate the SGD at the participant end,

$$\mathrm{yi} = \mathrm{yi} - \eta(\mathrm{gradient}(\mathrm{yi}))$$

Then the fedavg is repeated multiple times on each participant. The central serves aggregate all the SGD values, and these updated gradient values from each participant are shared to the central server. This makes global model but the centralized updates are not possible in this model which is solved by strat3.

$$x = \frac{1}{s}\sum yi$$

Strat3: SCAFFOLD - Stochastic Controlled Averaging for Federated Learning

This approach introduces a correction value of each gradient calculation of participants.

$$\mathrm{I} = \mathrm{yi} - \eta(\mathrm{gradient}(\mathrm{yi})) + (\mathrm{Co} - \mathrm{Coi})$$

Where (C_o-C_oi) is the correction term

3) The Applications of Federated Learning

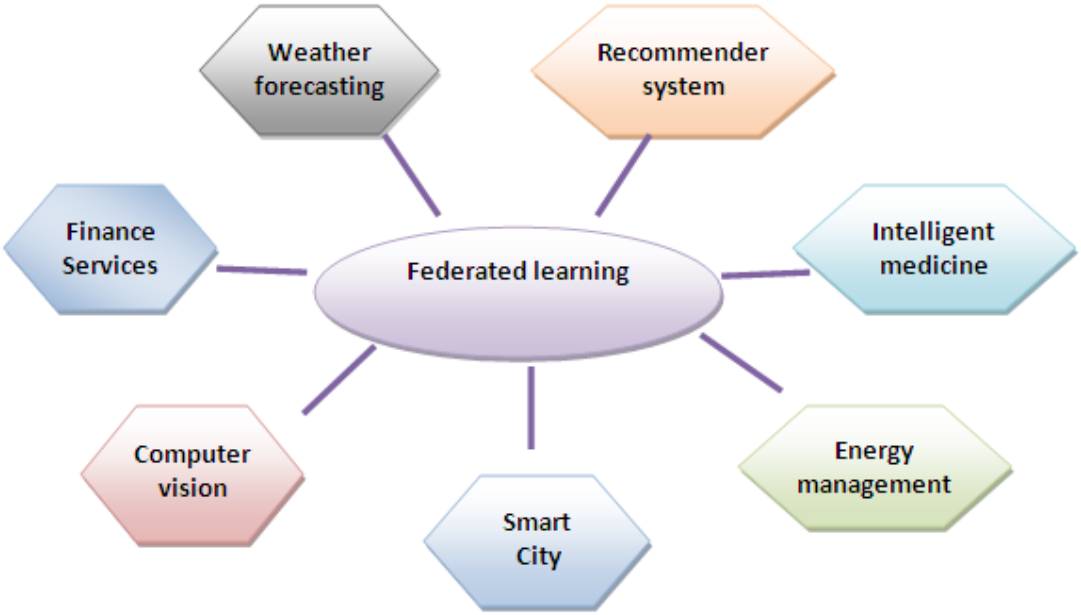

Figure 8: Applications of Federated Learning (FL)

In the big data environment, user's personal information is recorded in smart devices and edge servers so, data privacy shielding provides a guarantee for smart development of technologies. With the joint efforts of many scholars, FL has played an important position in the fields of diverse industries. And the present mainstream FL application areas are (illustrated in Figure 7) health care, finance, mobile applications, industry, computer vision and etc.

Healthcare: It protects sensitive data in the original source by gathering the data from various locations to diagnose a disease. Many researches claim that the future of digital wellbeing is with federated learning.

Autonomous Vehicles: FL can provide a safer self-driving with real time prediction of traffic and roads, real-time decision making and continual learning about the environment.

FedCV: it is unified library for FL to address the computer vision applications of image segmentation, preprocessing, classification and object detection. It is a federated learning framework to access various datasets and models through easy-to-use applications.

Recommendation system: the system learns the global model under restricted solitude requirements. Traditional recommender systems need to gather and analyze large amounts of user data to urge the service to the clients. The emergence of FL in this field provides a good framework to solve the problems includes social media, online platform, recommender and etc. The FL in recommendation system reduces the communication overhead across the federated networks.

Data privacy: By adopting federated learning, organizations can leverage the benefits of machine learning while preserving data security and privacy. It allows for collaborative model training on distributed data sources without the need to share raw data, making it an effective solution by using various data protection techniques as shown in Table 1 for various scenarios where data privacy and security are paramount concerns [19].

Table 1: The Different categories of attacks with their attack trait and protection techniques

Attack	Attack Trait	Protection Techniques
Byzantine attack [11]	Non-independent, identical distributed(non-IID) datasets	A credibility-based approach is used to defend against the attack
Byzantine attack [12]	Byzantine anti-conformance validation	Byzantine resistant secure block chained FL framework is introduced in which the validator does the parallel execution of heavy validation workflow
Poisoning attack [13]	Data poisoning	Detects and suppress the outliers for defending
Poisoning attack [14]	Data poisoning – Data source tags	The model trained from distributed vehicles is verified and stored block chain for privacy.
Poisoning attack [15]	Model poisoning	Developed a FL framework which uses consortium block chain technique

B. **Generative AI**

Generative AI models combines AI algorithms to represent and process various types of contents including text, image, audio, speech and synthetic data. These models are designed in a way to learn from vast amount of data and engender a new content that resembles original data. Traditional AI models are primarily used to analyze and predict the data, while generative AI is used to create a new data and contents similar to original data distribution. The key concept of GAI modeling is latent space, training data and generative architectures,

Latent space compressed representation of data with its essential features and removes unwanted features from the datasets. Training data are the base of learning and understands the underlying data patterns. Generative architectures are the building blocks of AI modeling. There are different models such as Variational Autoencoders (VAEs), auto-regressive models, flow-based models and Generative Adversarial Networks (GANs).

1) The Categorization of Generative AI Models

a) Variational Autoencoders (VAEs)

VAE has encoder-decoder architecture to map the input(x) into a latent space and they balance reconstruction accuracy along with regularization to generate new samples as described in Figure 9. It has two parts, encoder will encode the data from datasets and pass it to the bottleneck architecture and decoder uses latent space to regenerate the samples similar to dataset. The output (y) back propagates from the neural network in the form of loss function [11]. The model uses KL divergence as its loss function and aims to minimize the difference

between an intractable distribution and tractable distribution of dataset. The below steps elucidate the calculation of loss function using autoencoder,

Step 1: Consider a distribution Z and to generate the observation (O) from it and calculate

$$\text{P (Z/O) as } P(Z/O) = \frac{P(O/Z)P(Z)}{P(O)}$$

Step 2: Now calculate the intractable distribution P (O),

$$P(O) = \int P(O/Z)P(Z)dZ$$

Step 3: Once P(O) is calculated, now approximate intractable distribution P(O) to tractable distribution Q(O) by minimizing KL divergence loss,

$$minKL(Q(Z/O)||P(Z/O))$$

By simplifying the minimum function that is equivalwent to maximazation function,

$$E_{Q(Z/O)} \log P(O/Z) - KL(Q(Z/O)||P(Z))$$

The first term stands for reconstruction likelihood and other term represents that Q is similar to P.

Step 4: The loss function is calculated as,

$$loss = L(\text{o}, \hat{\text{o}}) + \sum (Q(Z/O)||P(Z))$$

Thus the total loss function has two terms, 1 term is reconstruction error and 2 term is KL divergence loss. VAEs have applications in assorted areas, including image generation, generation of realistic images, interactive exploration ,anomaly detection, and data compression.

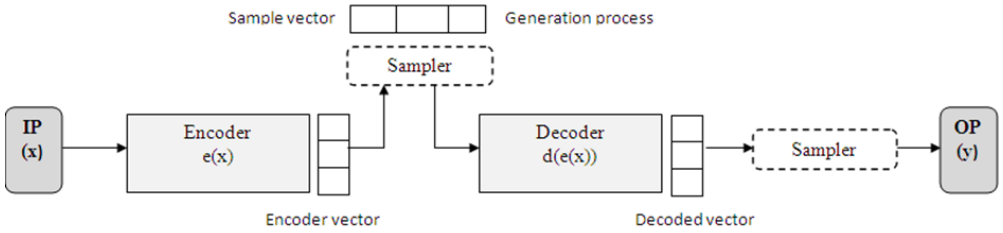

Figure 9: Variational Autoencoders (VAEs)

b) Generative Adversarial Networks (GANs)

GAN is deep learning architecture aims to generate a new, synthetic data that resemble known data distribution. It consists of three parts called Generative, Adversarial and Networks. In GAN's, there is a discriminator and a generator. The generator generates a sample of data to fool the discriminator as the discriminator finds which is fake and real image data. they both runs as competitors in the training phase to acquire better performance. GAN's have made contributions to video generation, image synthesis, style transfer, image to text synthesis, realistic simulation and etc.

c) Auto-Regressive Models

AR models forecasts current data behaviour from the past data. It is a linear model, where the current data is sum of past result multipled with numeric factor. Mathematically AR model represented as,

$$Y_t = c + \phi_t Y_{t-1} + \varepsilon_t$$

Where Y is time series variable, Y_{t-1} past data behaviour,c is numeric constant to reframe the lagged data, $\varepsilon_t$is the residual for predicting the time period t, ϕa numeric coefficient[20]. If a AR model has multiple lag then mathematically,

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-1} + \phi_3 Y_{t-1} + \varepsilon_t$$

Auto-regressive models are normally used in text creation, language modeling, and tune composition.

d) Flow-Based Models

FB models generates the image based on the sequence of inverse transformations between the input and output spaces. Normalizing flows were anticipated to solve the problems faced by GANs and VAE by using inverse transformation functions. Normalizing flows models provides the exact likelihood of the data and the models use negative log-likelihood as the loss function. FB models use inverse transformations to map the data point X with latent representation Z. The general representation of Flow based model is shown in Figure 10.

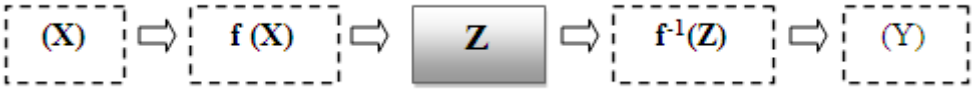

Figure 10: Representation of Flow-Based model

They allow generating the data and effectively estimating the density of data. They are widely used in anomaly detection, image generation, exact sampling, likelihood evaluation and density estimation.

2) The Applications of Generative AI

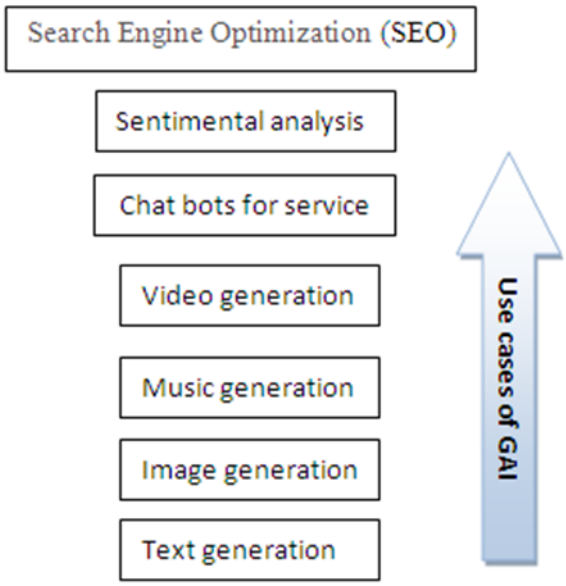

Figure 11: Use cases of Generative AI (GAI)

Table 2: Estimated market value of Artificial Intelligence (AI) in worldwide from 2020 to 2025

Market value in US dollars [16]	Year
12.5	2020
15.84	2021
20.82	2022
27.37	2023
35.99	2024
47.32	2025

The generative AI models are used widely (Figure 11) in various streams to create high quality content and patterns. Also, it helps in increasing market size of AI across the world as described in Table 2.

Image exploitation and creation: GAI creates realistic images and high-quality images from scratch that resembles real world objects. This model helps to translate one image to another and can assist artists to improve the quality the translated images.

Generation of text: These models create textual content including paragraphs, human like conversation, virtual assistants, chatbots, language translation and text summarization.

Generation of music compositions: Music is a significant component of many advertisements, and generative AI can be used to generate ad music for specific campaigns and services. This can increase the effectiveness of ads and drive more conversions in marketing fields. It also used in sound generation and virtual reality experiences.

Video Synthesis: Video generations are widely used in AI marketing to increase the engagement and sales of products. GAI models generate high quality video ads that can be used in various platforms and the models can mechanize video editing tasks, improve visual effects, create content in the film and entertainment industry. The future trends and development in GAI's to make advancements in GAI's architectures. The GAI models aim to improve the performance, efficiency, controllability of generative model and to integrate the generative models with other AI approaches including reinforcement and transfer learning process. AI generative models will continue to silhouette creativity and coerce innovation in unparalleled ways.

C. **Explainable AI**

AI has undergone much significant progress to solve different kinds of problems by using machine learning algorithm although their lack in transparency of decision making. Therefore, there need to come up with solutions to solve the mentioned challenge in the field of AI and sensitive domains. Explainable artificial intelligence (XAI) has been anticipated as a solution that can help to progress towards more translucent AI and it focuses on developing explainable techniques that allow users in trusting the decision making process (refer Figure 12 for the general framework of XAI model). Based on the survey, the need of XAI is explained in five perceptions,

1. Regulatory Perception, AI models are applied in our daily lives which could be resulting in intolerable decisions. So XAI is needed in explaining the intolerable decisions by which the end user can make a request to explain about the decision made by the algorithm and what is called the Right to explanation.
2. Scientific perception, it aims to develop a novel concept to address the given problem. Therefore, after creating a black box AI model, a novel model represents the basic knowledge rather than the data.
3. Industrial perception, user distrust about the black box AI models represents the challenges to the industry in applying multifaceted AI systems. XAI can alleviate between the interpretability and the performance to increase the development and deployment costs.
4. User and social perception, the probability to produce unfair decision is unease about black box system. Therefore, producing reason and explanation for the interpretability of the black-box AI systems will assist in increasing trust to the end users.
5. Model development perception, many reasons contribute to improper results including model overfitting, limited training and testing data, outliers and etc. so, the aim to use XAI is to understand, debug, and progress the black-box AI system to improve its robustness, increase safety and user trust, and minimize faulty actions, bias, wrongness, and unfairness [21].

The above perceptions highlight the most critical reasons why XAI is need in development process. The NIST [22] national institute of standard and technology states that four principles coerce the XAI,

- AI Systems deliver reasons and evidences for all acquired outputs.
- Systems provide explanation to the end user for understanding the decisions.
- The explanations correctly drive to the produced results.

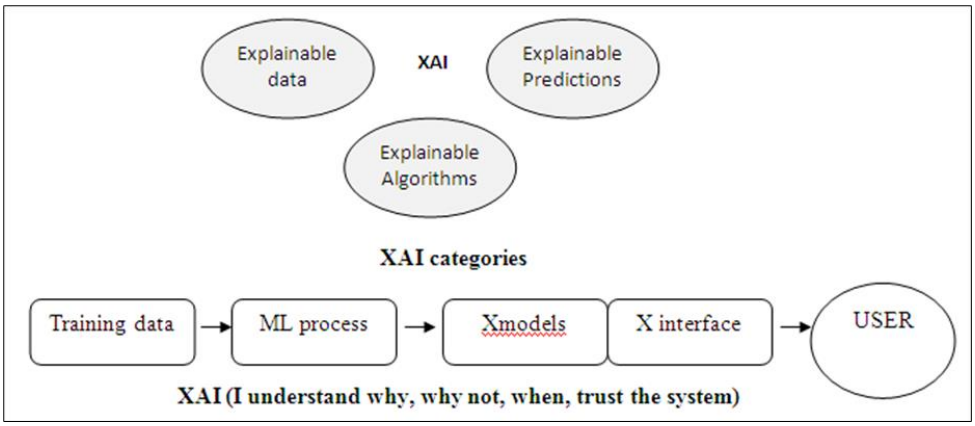

Figure 12: Framework of XAI model

1) The Explanation of XAI Principles

The NIST have explained five ways to apply the principles,

- Explain the subject of an algorithm to the end users.
- Explain the purpose of the algorithm to build the societal trust in an AI system.
- Satisfy the regulatory requirements in regulated industries.
- Benefit the algorithm/model owner.
- Assist with auxiliary system development.

2) The Applications of XAI Model

Healthcare: In diagnosing the disease, XAI model can enlighten the diagnosis process which helps doctors to explain about the diagnosis to patients. This creates trust between the patients and doctors for producing better results.

Manufacturing: This is important for enhanced machine-to-machine communication and understanding, which will help to build better situational consciousness between humans and machines.

Defense: AI used in military applications to explain about the decision made by the AI system which helps to mitigate ethical issues and challenges.

Autonomous vehicles: AI plays a major role in the automotive industry where explainability provides societal awareness in unexpected situations or accidents. XAI models are involved in safety critical decisions to reduce the risk of lives.

Loan approvals: Explainable artificial intelligence can be used to elucidate why particular loan was approved or denied for the user. This is imperative because it helps alleviate any potential ethical challenges by providing a better level of understanding between humans and machines.

Resume screening: Explainable artificial intelligence can be used to elucidate why particular resume was selected or not selected. This is imperative because it helps alleviate any potential ethical challenges by providing a better level of understanding between humans and HR system.

Fraud detection: if a transaction was suspicious, legitimate or flagged without a reason, then XAI can provide a solution to this kind of issue. XAI explains why transaction failed or flagged, which helps to identify the fraudulent transactions.

IV. FUTURE DIRECTIONS AND CONSTRAINTS

Positive AI technology future directions have enormous potential to alter various aspects of human existence, increase productivity, and address key global concerns.

Federated learning has the potential to change AI privacy. Federated learning protects individual data privacy while benefiting from a big, diversified dataset by allowing models to be trained locally on user devices and collecting knowledge without exchanging raw data. This approach could be extended in the future to collaborative settings involving different organizations, allowing secure knowledge sharing without jeopardizing critical information. Federated learning has implications in healthcare, as patient data can be kept local while medical models are collectively improved for better diagnosis and treatment. Furthermore, because it lowers the dangers associated with centralized data storage, it can encourage AI adoption in businesses with tight data protection requirements.

On the other hand, generative AI is poised to push the boundaries of creative content creation. Advances in generative models that can generate high-fidelity and diverse content, such as photos, music, movies, and even human-like prose, will be seen in the future. This has the potential to transform industries such as entertainment, art, and design by assisting human creators in developing new ideas and expressions. Furthermore, generative AI can enable interactive and immersive experiences, paving the way for novel applications in virtual and augmented reality. As technology advances, ethical considerations will become increasingly important in ensuring that AI-generated material complies with copyright rules and user consent.

Explainable AI (XAI) will support the responsible use of AI in crucial sectors such as healthcare and finance by allowing users to comprehend the rationale behind AI-driven decisions. AI integration with edge computing will enable quicker and more efficient real-time processing, allowing smart devices to operate autonomously and support applications such as smart cities and self-driving cars. AI-powered personalization will result in highly personalized user experiences, enhancing customer satisfaction and engagement across industries such as e-commerce, entertainment, and education.

A. **Challenges and Limitations**

In FL, there are various challenges to solve and to compromise the security of federated model as well as local model. In this section, we explained the core challenges in federated learning and the existing research works of each core challenges are briefly explained to understand the current enlargement of FL security protection mechanism. The implementation of FL is based on the challenges mention below,

- Communication efficiency in the network, the federated networks include a large of individual devices so the communication between the devices and network can be slower than the other methods. The following steps will achieve a communication efficient learning process, initially reduce the number of communication rounds and reduce the size of transmitted messages.
- Management of Heterogeneous system, the devices have different computation, communication capabilities in the federated network so the management of all the users and devices are biggest challenge in the FL.
- Statistical heterogeneity, Devices frequently make and collect data in a non-identically distributed manner across the network.
- Solitude concerns it deals to keep the raw data and information more secure. In addition, the privacy concern methods are adapted to the federated learning to secure the data of the participants. Recently, different privacy-preserving methods for machine learning have been researched. The solitude concerns of FL can be handled by typical privacy concern method like differential privacy, homomorphic encryption and secure multiparty computation (SMC).

In GAI metrics are there to assess the models, likelihood, inception score (IS), Frechet Inception Distance (FID) are used to assess the quality and evaluate the generated samples. Training Generative models can face more issues include overfitting of data, collapse of mode and finding the difference between exploration and exploitation. To generate a GAI model, should concern about ethical deliberation, transparency, fairness and deployment of models.

Similar to FL, XAI also faces several challenges in deployment and such challenges are

- Communication of data quality: The provided explanation of AI systems depends on the data used to build the model. The quality of data can be affected by the following reason like data bias, data incorrectness, data incompleteness [23]. So low quality of data doesn't only affect the output it leads to produce unfair decisions and degrades the explainability of the AI system.
- Sparsity of scrutiny: Validating the reason of an ML model may require examine process, which is a difficult task for the user to examine vast number of samples. Therefore, the number of samples for examination should be small as possible to reduce the sparsity of scrutiny [24]. The sparsity scrutiny is reduced by developing novel methods for identifying the meaningful subset of the dataset to interpret.
- Rules extraction: According to [25], there are three approaches for extraction of rules: a) the rules are extracted at the neuron level called as decomposition approach; b) Pedagogical approach, the rules are extracted based on the mapping of input to the output with its underlying structure, c) Eclectics approach, combination of decomposition and pedagogical approaches. Therefore, future work needs to be done to address the challenges connected with logic rule extraction and illustration for interpretable visual reasoning.
- Explaining the training process: Training ML models is a lengthy and tedious process because of large datasets is required to train the models. Therefore, researchers and practitioners are recommended in order to develop XAI models which support the online training process could help in developing better AI models and curtail time and resources
- Model innovation: By explaining AI models, can gain understanding of their internal structure which leads to existence of new models. Therefore, in future DL and new DL models are expected to complement each other which can helpful in bridging between the dense and translucent models.

V. CONCLUSION

The use of AI in the IT industry has resulted in a disruptive revolution, altering how businesses function and interact with technology. The ability of artificial intelligence (AI) to automate jobs, analyze massive volumes of data, and make intelligent conclusions has substantially improved operational efficiency, productivity, and decision-making processes. AI has become a vital tool for modern organizations, from intelligent chatbots and virtual assistants delivering seamless customer assistance to enhanced data analytics supporting data-driven insights. Furthermore, AI's applications span a wide range of industries, including healthcare, finance, manufacturing, and cybersecurity, among others, driving innovation and unlocking new opportunities. As AI advances, addressing issues such as ethical concerns, bias mitigation, and privacy concerns will be critical to ensuring responsible and sustainable AI adoption. The journey of the IT industry with AI is far from done, and the ongoing pursuit of AI's potential will pave the way for a future of limitless chances and improvements, influencing how businesses and society interact with technology for years to come.

REFERENCES

[1] J. Mageto, “Current and future trends of information technology and sustainability in logistics outsourcing,” Sustainability, vol. 14, no. 13, p. 7641, 2022.

[2] B. Li et al., “Trustworthy AI: From principles to practices,” ACM Comput. Surv., vol. 55, no. 9, pp. 1–46, 2023.

[3] H. Choung, P. David, and A. Ross, “Trust in AI and its role in the acceptance of AI technologies,” Int. J. Hum. Comput. Interact., pp. 1–13, 2022.

[4] “IDC: The premier global market intelligence firm,” IDC: The premier global market intelligence company. [Online]. Available: https://www.idc.com/. [Accessed: 30-Jul-2023].

[5] “China: GDP composition by industry 2022,” Statista. [Online]. Available: https://www.statista.com/statistics/1124008/china-composition-of-gdp-by-industry/. [Accessed: 30-Jul-2023].

[6] “India: IT-BPM industry share in GDP 2022,” Statista. [Online]. Available: https://www.statista.com/statistics/320776/contribution-of-indian-it-industry-to-india-s-gdp/. [Accessed: 30-Jul-2023].

[7] “Tech GDP as a percent of total GDP in the U.S. 2017-2022,” Statista. [Online]. Available: https://www.statista.com/statistics/1239480/united-states-leading-states-by-tech-contribution-to-gross-product/. [Accessed: 30-Jul-2023].

[8] Jetro.go.jp. [Online]. Available: https://www.jetro.go.jp/ext_images/en/invest/img/attractive_sectors/ict/ict_EN_202103.pdf. [Accessed: 30-Jul-2023].

[9] K. Aggarwal, “Has the future started? The current growth of artificial intelligence, machine learning, and deep learning,” Iraqi Journal for Computer Science and Mathematics, vol. 3, pp. 115–123, 2022.

[10] A. Heidari, N. J. Navimipour, and M. Unal, “Applications of ML/DL in the management of smart cities and societies based on new trends in information technologies: A systematic literature review,” Sustain. Cities Soc., vol. 85, no. 104089, p. 104089, 2022.

[11] X. Ma, Q. Jiang, M. Shojafar, M. Alazab, S. Kumar, and S. Kumari, “DisBezant: Secure and robust federated learning against byzantine attack in IoT-enabled MTS,” IEEE Trans. Intell. Transp. Syst., pp. 1–11, 2022.

[12] W. Li and S. Wang, “Federated meta-learning for spatial-temporal prediction,” Neural Comput. Appl., 2022.

[13] Y. Tian, W. Zhang, A. Simpson, Y. Liu, and Z. L. Jiang, “Defending against data poisoning attacks: From distributed learning to Federated learning,” Comput. J., vol. 66, no. 3, pp. 711–726, 2023.

[14] Y. Qi, M. Hossain, J. Nie, and X. Li, “Privacy-preserving block¬chain-based federated learning for traffic flow prediction. Future Generation Computer Systems-the,” International Journal of Escience, vol. 117, pp. 328–337, 2021.

[15] Y. Zhao, J. Chen, and J. Zhang, “Detecting and mitigating poisoning attacks in federated learning using generative adver¬sarial networks,” Concurrency and Computation: Practice and Experience, vol. 34, no. 7, 2022.

[16] C. Dilmegani, “Generative AI in marketing: Benefits & 7 use cases in 2023,” AIMultiple, 27-Feb-2023. [Online]. Available: https://research.aimultiple.com/generative-ai-in-marketing/. [Accessed: 30-Jul-2023].

[17] X. Liu, H. Li, G. Xu, R. Lu, and M. He, “Adaptive privacy-preserving federated learning,” Peer Peer Netw. Appl., vol. 13, no. 6, pp. 2356–2366, 2020.

[18] J. Wen, Z. Zhang, Y. Lan, Z. Cui, J. Cai, and W. Zhang, “A survey on federated learning: challenges and applications,” Int. J. Mach. Learn. Cybern., vol. 14, no. 2, pp. 513–535, 2023.

[19] N. Rieke et al., “The future of digital health with federated learning,” NPJ Digit. Med., vol. 3, no. 1, p. 119, 2020.

[20] V. Mehandzhiyski, “What is an Autoregressive Model?,” 365 Data Science, 06-Mar-2020. [Online]. Available: https://365datascience.com/tutorials/time-series-analysis-tutorials/autoregressive-model/. [Accessed: 30-Jul-2023].

[21] W. Saeed and C. Omlin, “Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities,” Knowl. Based Syst., vol. 263, no. 110273, p. 110273, 2023.

[22] “National institute of standards and technology | NIST.”

[23] E. Reiter, “Natural language generation challenges for explainable AI,” in Proceedings of the 1st Workshop on Interactive Natural Language Technology for Explainable Artificial Intelligence (NL4XAI 2019), 2019.

[24] G. Dao and M. Lee, Demystifying deep neural networks through interpretation: A survey(2020). 2020.

[25] C. He, M. Ma, and P. Wang, “Extract interpretability-accuracy balanced rules from artificial neural networks: A review,” Neurocomputing, vol. 387, pp. 346–358, 2020.