

www.ijprems.com editor@ijprems.com

Vol. 04, Issue 06, June 2024, pp: 1271-1279

A BRIEF IDEA OF AI AND ITS APPLICATION TO VARIOUS FIELDS

Dr. Rajesh Kumar¹, Ms. Vandana Sharma²

¹HOD-BCA Department of Information, Communication & Technology Tecnia Institute of Advanced Studies is a private college affiliated with Guru Gobind Singh Indraprastha University and located in Madhuban Chowk Rohini, Delhi, India.

²Assistant Professor Department of Information, Communication & Technology Tecnia Institute of Advanced Studies is a private college affiliated with Guru Gobind Singh Indraprastha University and located in Madhuban Chowk Rohini, Delhi, India.

DOI: https://www.doi.org/10.58257/IJPREMS34947

ABSTRACT

An introduction to artificial intelligence (AI) is provided in this paper. Artificial entities, which are typically thought to be computers, display artificial intelligence. Artificial Intelligence (AI) systems are now widely used in fields such as economics, engineering, medicine, and the military. They are also integrated into many popular home computer software applications, classic strategy games like computer chess, and other video games.

We tried to explain the brief ideas of AI and its application to various fields. It cleared the concept of computational and conventional categories. It includes various advanced systems such as Neural Network, Fuzzy Systems and Evolutionary computation. AI is used in typical problems such as Pattern recognition, Natural language processing and more. This system is working throughout the world as an artificial brain.

Artificial intelligence (AI) research has found ways to have computers do some of the procedures that make up intelligence, but not others. When a goal just requires current, well-understood mechanisms, computer programs can accomplish quite astounding results. These programs ought to be regarded as "somewhat intelligent." It has something to do with the comparable task of comprehending human intelligence through computer use.

By studying our own processes or those of others, we can gain some insight into the process of programming machines to solve issues. However, rather than researching humans or animals, the majority of AI research focuses on the issues the outside world poses to intelligence. Researchers in AI are free to employ techniques that are not seen in humans or that require a lot more processing power than is humanly possible. We spoke about the requirements for judging a machine to be intelligent. We contended that you should consider the machine intelligent if it could effectively pass for a human to an informed observer.

A basic definition of intelligence is the ability to solve problems, both real and imagined, in a way that yields a useful end result. Intelligence also involves the capacity to discover new problems, both real and imagined, which allows one to build on existing knowledge.

Keywords: GPT-1, OpenAI, GPT-3, Google Bard, GPT-4, Natural Language Processing, Machine Learning, Artificial Intelligence Robotics, Autonomous Technology, AI Ethics, Human-Computer Interaction

1. INTRODUCTION ARTIFICIAL INTELLIGENCE (AI)

The field known as "artificial intelligence" studies how computers can learn to solve complicated problems in a way that mimics human intelligence. In most cases, this entails taking cues from human intellect and implementing them as computer-friendly algorithms. The behaviour of artificial intelligence is influenced by the requirements that are set, which might lead to more or less flexibility or efficiency in the method used.

While computer science is often thought of as the birthplace of artificial intelligence, the concept actually has deep ties to many other disciplines, including but not limited to mathematics, philosophy, biology, psychology, and cognition. We can make more headway in our effort to create an intelligent artificial creature if we can integrate information from all these domains.

The public's perception of artificial intelligence is largely based on its use in robotics research and development, but the technology has found its most practical use as an integrated part of software in domains like data mining, physical science, economics, and finance and economics. The goal of artificial intelligence (AI) in robotics is to simulate human thought using computer models. Creating software that mimics human behaviour is insufficient. You aim to create an algorithm that mimics human performance. Computer scientists also face challenges while attempting to develop artificial intelligence (AI) systems capable of comprehending complex, high-level languages.

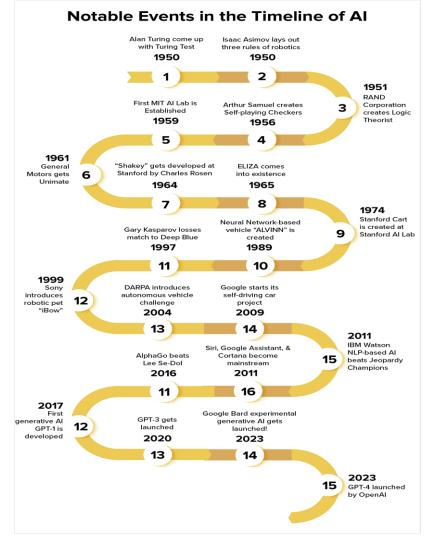


editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

Vol. 04, Issue 06, June 2024, pp: 1271-1279

2. AI HISTORY



Here is now explain AI journey 1950 to Explanation is own words

S.No	Year	Research/Invent	Explanation
1	1950	Turing Test	Alan Turing introduces the Turing Test as a means of assessing a machine's capacity to demonstrate intelligent behaviour that is on par with, or indistinguishable from, that of a person.
2	1950	Three Rules of Robotics	The Three Laws of Robotics, proposed by Isaac Asimov, are meant to guarantee that robots act ethically.
3	1951	Logic Theorist	The RAND Corporation creates the Logic Theorist, considered the first artificial intelligence program, which could mimic human problem-solving skills.
4	1956	Self-Playing Checkers	An early achievement in machine learning is the algorithm that Arthur Samuel developed that can play checkers against a human opponent.
5		MIT AI Lab	The AI Laboratory, a major hub for AI research, is established at the Massachusetts Institute of Technology (MIT).
6	1961	The Unimate	Unimate is the first industrial robot, and General Motors uses it to help out on the assembly line.
7	1964	Shakey	Charles Rosen of Stanford University creates "Shakey," the first



INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062 Impact Factor:

5.725

www.ijprems.com editor@ijprems.com

Vol. 04, Issue 06, June 2024, pp: 1271-1279

			mobile robot with general-purpose capabilities and the ability to reason about its own behaviours.
8	1965	ELIZA	Joseph Weizenbaum develops ELIZA, a pioneering software in natural language processing that could mimic human speech.
09	1972	Stanford Cart	The Stanford Cart was developed by Stanford University's AI Lab as a preliminary test of autonomous vehicles.
10	1989	ALVINN	The ALVINN (Autonomous Land Vehicle In a Neural Network) is an autonomous vehicle that relies on neural networks and was developed by Carnegie Mellon University.
11	1991	Deep Blue vs. Kasparov	World chess champion Garry Kasparov is defeated by IBM's Deep Blue computer, demonstrating the strength of artificial intelligence in difficult strategic games.
12	1991	Sony's AIBO	A major milestone in consumer robotics and artificial intelligence is the debut of AIBO by Sony, a robotic pet with interactive features.
13	2004	DARPA Grand Challenge	Participants in DARPA's Grand Challenge work to improve autonomous vehicle technology; the goal is to create vehicles that can drive themselves across challenging off-road terrain.
14	2009	Google's Self-Driving Car	Much progress has been made in the development and public awareness of autonomous vehicle technology with the launch of Google's self-driving car project.
15	2011	IBM Watson	By defeating human champions in Jeopardy!, IBM Watson demonstrates the superior data retrieval and natural language processing skills of artificial intelligence.
16	2011	Siri, Google Assistant, Cortana	The introduction of AI-powered virtual assistants like Cortana, Google Assistant, and Siri has revolutionized the way people engage with technology in their daily lives.
17	2016	AlphaGo vs. Lee Se-Dol	AlphaGo, developed by Google's DeepMind, beats Go champion Lee Se-Dol, proving that AI can outsmart humans in abstract, complicated strategy games.
18	2017	GPT-1	With the release of GPT-1, the first Generative Pre-trained Transformer, OpenAI has made a giant leap forward in artificial intelligence's capacity to comprehend and produce language that sounds natural.
19	2020	GPT-3	With the introduction of GPT-3, an advanced language model, OpenAI revolutionizes AI applications with its ability to generate coherent, human-like prose and execute a broad variety of language tasks.
20	2023	Google Bard	To boost creative and conversational apps with enhanced language comprehension and generating skills, Google develops Bard, an experimental generative AI.
21	2023	GPT-4	In keeping with the ongoing development of artificial intelligence, OpenAI has released GPT-4, which improves generative language models via increased capability, accuracy, and application variety.



editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

Vol. 04, Issue 06, June 2024, pp: 1271-1279

e-ISSN : 2583-1062 Impact Factor: 5.725

3. GOALS OF AI

Several distinct sub-problems have been identified within the larger issue of intelligence simulation or creation. These are the characteristics and skills that scientists hope an AI system will have. Most of the focus has been on the characteristics listed below.

Reasoning, deducing, and solving problems:

A "combinatorial explosion" occurs when the amount of memory or computer time required becomes astronomical as the problem gets above a certain size, meaning that most of these algorithms can require enormous computing resources for tough situations. One of the main goals of artificial intelligence research is the development of algorithms that solve problems more efficiently.

Instead of the deliberate, sequential reasoning that early artificial intelligence studies may mimic, humans typically resort to quick, intuitive decisions when faced with a challenge. Artificial intelligence has come a long way in the area of "sub-symbolic" problem solving, with methods such as embodied agent approaches, neural net research (which aims to replicate the brain structures that give rise to this skill), and statistical AI (which mimics the probabilistic nature of human guessing) all making strides in this direction.

The depiction of knowledge: -

Artificial intelligence studies revolve around knowledge representation and engineering. Extensive world knowledge will be necessary for many of the problems that machines are supposed to handle. Objects, attributes, classes, and relationships among objects; contexts, occurrences, states, and time; effects and causes; knowledge about knowledge (i.e., what we know about other people's knowledge); and a plethora of additional, less explored areas are all things that AI must represent. An ontology is a representation of "what exists"; it is the collection of things, relations, ideas, etc. that the computer is aware of. The highest level of generality is attained by higher ontologies, which endeavour to lay the groundwork for all other forms of knowledge.

Planning:-

Agents with intelligence need to be able to plan ahead and complete tasks. In order to make decisions that optimise the utility (or "value") of their options, they require a means of seeing into the future.

The agent in classical planning issues can know for sure what its actions will do to the world and can operate as if no one else is involved. The agent must reason under uncertainty if it is not the sole actor; otherwise, it will need to check in with itself at regular intervals to see if reality is conforming to its predictions and adjust its strategy accordingly.

Natural language processing:-

Machines can now read and comprehend human languages thanks to natural language processing. To acquire knowledge directly from human-written sources, such Internet writings, and to enable natural language user interfaces, one needs a natural language processing system that is powerful enough. Machine translation and information retrieval (sometimes known as text mining) are two simple uses of natural language processing. In the field of natural language processing, semantic indexing is a popular tool for meaning extraction. Indexing massive amounts of user input abstracts is now significantly more efficient thanks to increases in processing speeds and decreases in the cost of data storage.

In the field of natural language processing, semantic indexing is a popular tool for meaning extraction. Indexing massive amounts of user input abstracts is now significantly more efficient thanks to increases in processing speeds and decreases in the cost of data storage.

Dynamics and control:

Robotics and artificial intelligence (AI) go hand in hand in many ways. To carry out activities such as travelling and object manipulation, robots require intelligence. The sub problems of localization, mapping, and path or motion planning are crucial to these activities. Making a map requires research into the surrounding area and the recording of that information. One example is compliant motion, where the robot manipulates an object by physically contacting it.

The View: -

One definition of machine perception is the capacity to infer features of the environment from data collected by various sensors (even more exotic ones) like sonar, cameras, microphones, and microphones. Vision in a computer system is the capacity to interpret images. Recognising faces, objects, and spoken language are a few of the chosen subproblems.



editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

Vol. 04, Issue 06, June 2024, pp: 1271-1279

Interpersonal competence:-

The field devoted to the research and creation of technology with the ability to detect, understand, process, and mimic human emotions is known as affective computing. The roots of this multidisciplinary area can be found in early philosophical investigations into emotion, but nowadays it encompasses computer technology, psychology, and cognitive science. One reason to do the research is to find a way to mimic empathy. Artificial intelligence systems should be able to detect when people are feeling sad or angry and adjust their actions accordingly.

The ability to feel and interact with others serves a dual purpose for intelligent beings. To start, it needs peoplereading skills to know what other people are thinking and feeling so it can anticipate their moves. (This calls for knowledge of decision theory and game theory, abilities to model human emotions, and the capacity to perceive emotions.) Additionally, an AI may seek to emulate human empathy by creating the illusion of emotional sensitivity in its interactions with humans, even when it lacks the capacity to truly feel such emotions.

General intelligence:-

Eventually, most researchers hope that their work will be integrated into a machine that has general intelligence, also called strong AI. This machine will possess all the capabilities that humans are good at plus some. There are some who think it could be necessary to incorporate human-like characteristics into such a project, such as an artificial brain or artificial consciousness. Solving several of the aforementioned challenges might necessitate the application of general intelligence. Machine translation is just one example of a simple, targeted activity that involves natural language processing (NLP), reasoning, knowledge, and social intelligence to accurately recreate the author's intent. One example of a problem that is deemed "AI-complete" is machine translation. You have to fix every single problem before you can fix this one.

4. AI CLASSIFICATIONS

As far as AI is concerned, there are essentially two camps:

Classical Automaton: -

Formalism and statistical analysis define traditional AI, which primarily employs what is presently known as machine learning. This goes by a few other names: Good Old Fashioned Artificial Intelligence (GOFAI), symbolic AI, logical AI, and tidy AI.

Techniques involve:

Professional systems: use their thinking talents to determine a result. A significant quantity of known information may be processed and used to draw judgements by an expert system.

- Relying on example
- ✤ Bayesian models
- Bayesian models
- Behavioral AI: a hands-on approach to constructing AI systems using a modular design.

Smart Machines (SM):-

Iterative development or learning is a key component of computational intelligence, as shown in processes such as parameter tweaking in connectionist systems. Soft computing, non-symbolic AI, and learning based on empirical data are all terms that describe the same thing.

Here are several methods:

- Systems equipped with exceptional pattern recognition capabilities: neural networks.
- Control systems for both commercial and consumer goods have made extensive use of fuzzy logic, which refers to methods for reasoning under uncertainty.

In evolutionary computing, ideas borrowed from biology, such as natural selection, mutation, and population dynamics, are used to find better and better solutions to problems. The two main branches of these approaches are swarm intelligence (e.g., ant algorithms) and evolutionary algorithms (e.g., genetic algorithms).

Common problems that AI methods attempt to solve:-

Pattern identification system

- Automatic OCR
- Recognition of handwriting
- Voice detection
- Facial identification
- Chatbots, translation, and natural language processing



editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

2583-1062 Impact Factor:

5.725

e-ISSN:

Vol. 04, Issue 06, June 2024, pp: 1271-1279

- The field of robotics and non-linear control
- ✤ Virtual reality, image processing, computer vision
- Strategic planning and game theory

Extra domains where AI techniques are put into practice:-

✤ Automating tasks:

Machines, control systems, and information technology are the building blocks of automation, which maximises output in manufacturing and service provision. The right motivation to use automation is to achieve consistent quality standards and/or boost productivity beyond what is achievable with the present levels of human labour. This will allow for economies of scale. Automating formerly manual processes substantially reduces reliance on human senses and cognition while simultaneously boosting throughput, velocity, and consistency.

***** Studying cybernetics:

- When it comes to the dynamic aspect of the system being structured, cybernetics places a particular focus, making it similar to the science of organisation in certain respects. One organisation that meets the criteria for cybernetic investigation is the human brain. It is typical of many government agencies and huge organisations, and it contains all the features of feedback and storage, etc.
- One branch of AI is cybernetics, which seeks to demonstrate intelligent activity in artificially constructed systems.
- ✤ Hybrid intelligent system:-
- An innovative way to build computationally intelligent systems is to hybridise different intelligent systems. These systems can include algorithms like evolutionary computation, swarm intelligence, bacterial foraging, rough sets, approximate reasoning, fuzzy inference systems, and artificial neural networks. A plethora of new, superior forms of intelligence, dubbed Hybrid Intelligence, have emerged in recent years as a result of the merging or hybridization of various learning and adaptation techniques in an effort to overcome individual limitations and achieve synergistic effects.

✤ A smart agent:

Intelligent agents (IAs) in the field of artificial intelligence are self-governing entities that monitor their surroundings via sensors and manipulate it through actuators (i.e., they are agents) in order to accomplish predetermined objectives.

Control that is intelligent

A new field called intelligent control, sometimes known as self-organizing/learning control, is developing to address these issues. It is more based on actual experience than on theoretical models. Integrating AI, ST, and OR into a single framework, Intelligent Control is a promising new field. It incorporates new information into its performance by use of different learning schemas, which, in order to be put into practice, need to show signs of fast learning convergence, stability over time, resilience in the face of parameter changes, and internal and external disruptions.

* Reasoning on the fly: -

 Computer programmes that can reason entirely (or almost entirely) on their own are a product of research into automated reasoning. Though it is mostly associated with AI, automated reasoning does have ties to theoretical CS and even philosophy.

Mining data:

An interdisciplinary branch of computer science, data mining is the analytical phase of the "Knowledge Discovery in Databases" (KDD) process. It is the computational method of finding patterns in big data sets using techniques at the crossroads of AI, ML, statistics, and database systems. Data mining's overarching purpose is to extract useful information from datasets by transforming them into a comprehensible framework.

Behaviour-based robotics:-

- ✤ A subfield of robotics known as behaviour-based robotics brings together experts in cognitive science, engineering, and artificial intelligence (AI). On one hand, it aims to:
- The goal is to provide techniques for controlling many types of artificial systems, including real robots, virtual ones, and intelligent software agents.
- The goal is to simulate and comprehend biological systems, usually include creatures as diverse as humans and insects, using robots. Robots that can think.



INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com editor@ijprems.com

Robots for development: -

- An approach to building autonomous robot minds, Developmental Robotics (DevRob) draws on analogies from neuroscience and psychology of child development. It is also known as epigenetic robotics.
- A developmental programme is software that mimics the genome's operations in order to build a robot's cognitive capacities.

Robotics based on evolution: -

• Evolving controllers for robots using evolutionary computation is the main goal of evolutionary robotics (ER).

Chatbot:

- A chatterbot is an example of a conversational agent, a software programme that can mimic human speech and body language in order to carry on natural-sounding conversations with humans.
- A client-side script or standalone application that establishes an Internet Relay Chat connection and masquerades as another user to other IRC users is known as an Internet Relay Chat bot.

Knowledge Representation:-

- Information representation (KR) is a subfield of artificial intelligence that studies how to symbolically represent information in order to make inferences from those pieces and ultimately create new knowledge.
- One way to make the KR stand on its own is to separate it from the semantic network or other underlying knowledge model or KBS.

5. AI'S REAL-WORLD USES

Numerous domains have made use of artificial intelligence, such as legal, medical, stock trading, robot control, scientific discoveries, and toys.

Medical centres and hospitals:-

Artificial intelligence systems may help medical clinics with tasks such as staff rotation, medical information provision, and bed schedule organisation.

As an example, EMR software's Concept Processing uses artificial neural networks as a clinical decision support system for medical diagnosis.

- intelligence Artificial has the ability carry out additional medical duties such as: • to Computer-aided interpretation of medical imaging. These methods are useful for analysing digital pictures (such as those from computed tomography) for normal features and for identifying abnormalities in more noticeable areas, including potential illnesses. The identification of tumours is an example of a common use.
- Assessment of cardiac sonography.

Industrial strength

Many industries have begun to use robots. They are often assigned tasks that are deemed hazardous to human health. In tasks where human error is possible due to distraction or monotony, as well as in tasks that people find demeaning, robots have shown to be an effective alternative.

Engaging in Gaming

- During the height of the digital revolution, this flourished, exposing people—particularly kids—to the possibility of a future filled with AI.
- Also, for a few hundred bucks, you can get a computer that can play chess at a master's level. Even while they include some artificial intelligence, they mostly outperform humans by analyzing hundreds of thousands of locations via brute force processing.
- > One of the finest places to get machines and play different games is on the internet.

Voice Recognition

It wasn't until the 1990s that computer voice recognition was usable for some restricted tasks. As a result, United Airlines has ditched its flight information keyboard tree in favor of one that can recognize flight numbers and city names using voice recognition. The convenience is great. Still, most people find that using a keyboard and mouse is more convenient, even if certain computers can be instructed verbally.

Deciphering Natural Language:

A computer needs more than just a string of text. Neither is sentence parsing sufficient. At the moment, this is only achievable for a small subset of domains since the computer has to be taught about the subject matter of the text.



www.ijprems.com editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

Vol. 04, Issue 06, June 2024, pp: 1271-1279

e-ISSN : 2583-1062 Impact Factor: 5.725

Computer Vision

Although our eyes and computers only get two-dimensional images, the things in the real world are really threedimensional. Full computer vision needs partly three-dimensional data, not only a collection of two-dimensional images, while some practical applications may function in just two dimensions. Direct representations of threedimensional information are currently restricted and clearly inferior to what people utilize.

Expert Systems

In order to encapsulate domain experts' knowledge in a computer program for task execution, a "knowledge engineer" conducts interviews with them. To what extent this succeeds is dependent on whether or not the necessary cognitive processes are within the capabilities of current AI. In 1974, MYCIN was one of the first expert systems that could detect and recommend therapies for blood-borne bacterial illnesses. Following the observance of its constraints, it outperformed medical students and actual physicians.

Classification using Heuristics

Using several sources of data to classify data into predetermined categories is one of the most practical types of expert systems available today in the field of artificial intelligence. The recommendation to approve or reject a credit card transaction is one such example. It is possible to find out details about the person using the credit card, their payment history, the product they are purchasing, and even the store they are purchasing it from (such as whether or not this store has a history of credit card fraud).

6. APPROACH TOWARDS

ARCHITECTURE OF FUTURE AI

- Advancements in some areas, like voice recognition, will keep getting better over the next decade, eventually reaching human levels.
- In ten years, AI will have some common sense, be able to traverse an unprepared environment (not flawlessly), and converse with people in unstructured English via voice or text.
- Parts of the animal brain, including the human brain, will be replicated in silicon. First hippocampal tests in rats show that this is feasible. Both CCortex and IBM Blue Brain are significant endeavors with the goal of simulating the human brain.
- Practical applications of computer recreations of human intelligence, including cognition, perception, rehearsal learning, and learning via recurrent practice, will proliferate.
- > In order for robots to build AI that is meaningful, they will need to mimic human awareness in some way.
- > Without the capacity for self-awareness and sentience, systems can only ever be very fragile.
- Really beneficial and strong assistants will continue to be an aspirational objective in the absence of these distinctively human traits. Undoubtedly, ever-increasing functional leaps will be made possible by advancements in hardware, storage, and parallel processing architectures.
- Moving ahead, systems will be required to prove, with certainty, that they have self-awareness, linguistic abilities, and surface, superficial, and deep understanding of their environment and their place in it.
- > Nonetheless, research on AI is still in its early stages.
- \succ Nevertheless, this field ought to see tremendous progress in the early years of the new millennium.

7. CONCLUSION

Finally, we say that you should definitely consider the machine intelligent if it could fool a knowledgeable observer into thinking it's human. Not only are AI systems integrated into several popular home computer software programs, conventional strategy games, and more, but are also routinely used in fields like economics, medicine, engineering, and the military. The field of artificial intelligence is both fascinating and lucrative. The study of how to program computers to act intelligently is known as artificial intelligence (AI).

The new definition of artificial intelligence is that it is the study of intelligent behavior's underlying processes by building and testing artifacts that try to mimic those mechanisms. This leads us to believe that it functions similarly to a human brain, but with much superior artificial intelligence.

8. REFERENCE

- [1] Zoph, B., & Le, Q. V. (2017). Neural architecture search with reinforcement learning. arXiv preprint arXiv:1611.01578.
- [2] Liu, H., Simonyan, K., & Yang, Y. (2018). DARTS: Differentiable architecture search. arXiv preprint arXiv:1806.09055.



editor@ijprems.com

e-ISSN: **INTERNATIONAL JOURNAL OF PROGRESSIVE** 2583-1062 **RESEARCH IN ENGINEERING MANAGEMENT** AND SCIENCE (IJPREMS) Impact

Vol. 04, Issue 06, June 2024, pp: 1271-1279

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on Explainable Artificial [3] Intelligence (XAI). IEEE Access, 6, 52138-52160.

- [4] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1135-1144).
- [5] Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. IEEE Transactions on knowledge and data engineering, 22(10), 1345-1359.
- [6] Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? In Advances in neural information processing systems (pp. 3320-3328).
- Levine, S., Pastor, P., Krizhevsky, A., & Quillen, D. (2016). Learning hand-eye coordination for robotic [7] grasping with deep learning and large-scale data collection. arXiv preprint arXiv:1603.02199.
- [8] Schulman, J., Levine, S., Moritz, P., Jordan, M., & Abbeel, P. (2015). Trust region policy optimization. In Proceedings of the 32nd International Conference on Machine Learning (ICML-15) (pp. 1889-1897)
- [9] Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2018). Deep learning for healthcare: review, opportunities and challenges. Briefings in bioinformatics, 19(6), 1236-1246.
- [10] Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., ... & Ng, P. (2018). Scalable and accurate deep learning with electronic health records. npj Digital Medicine, 1(1), 1-10
- [11] Goodfellow, I. J., Shlens, J., & Szegedy, C. (2015). Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572.
- [12] Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., & Fergus, R. (2013). Intriguing properties of neural networks. arXiv preprint arXiv:1312.6199.
- [13] Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389-399.
- [14] Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., ... & Gebru, T. (2019). Model cards for model reporting. In Proceedings of the conference on fairness, accountability, and transparency (pp. 220-229)
- [15] Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2016). Federated learning: Strategies for improving communication efficiency. arXiv preprint arXiv:1610.05492.
- [16] McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2016). Communication-efficient learning of deep networks from decentralized data. arXiv preprint arXiv:1602.05629.
- [17] https://ieeexplore.ieee.org/abstract/document/9521302
- [18] https://pubmed.ncbi.nlm.nih.gov/33596342/
- [19] https://arxiv.org/abs/2003.09593
- [20] https://www.sciencedirect.com/science/article/pii/S2213158220302107
- [21] https://arxiv.org/abs/2006.00749
- [22] https://arxiv.org/abs/2009.00951
- [23] https://www.nature.com/articles/d41586-019-02919-8