



The Future and Risques of Artificial Intelligence

Shin Ah Son[†]

Received date: 29-Mar-2022, Manuscript No. NPY-22-57989; **Editor assigned date:** 31-Mar-2022, PreQC No. NPY-22-57989(PQ); **Reviewed date:** 11-Apr-2022, QC No NPY-22-57989; **Revised date:** 21-Apr-2022, Manuscript No. NPY-22-57989(R); **Published date:** 28-Apr-2022, DOI: 10.37532/1758-2008.2022.12(4).636

Description

Artificial Intelligence (AI) refers to machine intelligence rather than natural intelligence created by animals such as humans. The study of “intelligent agents,” or systems that understand their surroundings and take actions that maximize their chances of achieving their goals, as defined by leading AI textbooks. Leading AI researchers, on the other hand, reject this definition, which utilizes the word “artificial intelligence” to refer to robots that mimic “cognitive” functions that humans associate with the human mind such as “learning” and “problem solving.”

The AI effect is a phenomenon that occurs when machines become more capable, and tasks that were once thought to require “intelligence” are increasingly removed from the AI idea. Optical character recognition, for example, is frequently overlooked in AI debates, despite the fact that it is a widely used technology. Since its inception as a field of research in 1956, Artificial Intelligence (AI) has gone through several waves of enthusiasm, disappointment, and funding loss followed by new approaches, success, and renewed investment. Since its inception, AI research has experimented with and rejected a variety of approaches, including brain mimicking, human problem solving modeling, formal logic, enormous knowledge libraries, and animal behaviour imitation.

Highly mathematical statistical machine learning dominated the subject in the first decades of the twenty-first century, and this technique has prov-

en highly successful, helping to tackle many tough problems in industry and academics. The many sub-fields of AI research are based on specific aims and the application of certain techniques. Traditional AI research goals include reasoning, knowledge representation, planning, learning, natural language processing, sensing, and the ability to move and manipulate objects. General intelligence is one of the field’s long-term aims (the capacity to solve any problem).

To deal with these challenges, AI researchers have adapted and incorporated a variety of problem-solving tools, such as search and mathematical optimization, formal logic, artificial neural networks, and statistics, probability, and economics methodologies. AI incorporates computer science, psychology, linguistics, philosophy, and a variety of other fields. Human intelligence “can be so clearly described that a machine can be constructed to imitate it,” according to the field’s founders. This sparks philosophical debates about the mind and the ethics of building artificial intelligence that is human-like. Since antiquity, myth, fiction, and philosophy have all attempted to address these challenges. AI, with its vast potential and power, has also been suggested as an existential threat to humans in science fiction and futurology.

Future

A hypothetical agent with intelligence considerably above that of the brightest and most gifted human mind is known as a super intelligence, hyper intelligence, or superhuman intelligence. The form or degree of intelligence possessed by such an agent is

Department of Thoracic and Cardiovascular Surgery, Kyungpook National University Hospital, Daegu, Korea

[†]Author for Correspondence: Shin Ah Son, Department of Thoracic and Cardiovascular Surgery, Kyungpook National University Hospital, Daegu, Korea email: sinaah.as@gmail.com.

also referred to as super intelligence. If artificial general intelligence research yielded sufficiently clever software, it might be able to self-regulate and improve. The enhanced software would become increasingly better at improving itself, resulting in recursive self-improvement. In an intelligence explosion, its intelligence would skyrocket, surpassing humans by a factor of ten. A planning intelligent agent builds a model of the world, predicts how their actions will affect it, and makes decisions that maximize the utility (or “value”) of the available possibilities. In classical planning issues, the agent can assume it is the sole system acting in the world, which gives it confidence in the results of its actions. However, if the agent isn’t the lone player, it must reason in the face of uncertainty and constantly re-evaluate and alter its surroundings. Multi-agent planning combines the collaboration and competition of several agents to attain a certain goal. Emergent behaviour is exploited by both evolutionary algorithms and swarm intelligence.

Vernor Vinge, a science fiction writer, coined the term “singularity” to describe this scenario. The technological singularity is an occurrence beyond which events are unpredictable or even unfathomable since it is difficult or impossible to know the limits of intelligence or the capabilities of super intelligent computers. Objects, attributes, categories, and interactions between objects; situations, events, states, and time; causes and consequences; knowledge about knowledge; default thinking; and other domains have all been represented by AI research. The breadth of commonsense knowledge and the sub-symbolic structure of most commonsense knowledge are two of AI’s most difficult difficulties. Content-based indexing and retrieval, scene interpretation, clinical decision support, knowledge discovery, and other domains all use formal knowledge representations.

Risks

Smart spyware, face recognition, and voice recognition allow widespread surveillance; such surveillance allows machine learning to classify

potential state enemies and prevent them from hiding; recommendation systems can precisely target propaganda and misinformation for maximum effect; deep fakes aid in the production of misinformation; advanced AI can make centralized decision markets. Other forms of weaponized AI, such as advanced digital warfare and lethal autonomous weaponry, may be used by terrorists, criminals, and rogue nations. Over fifty countries were said to be working on battlefield robots by 2015.

After learning from real-world data, AI programmes may become prejudiced. Because it is learned by the software rather than introduced by the system designers, programmers are frequently ignorant of the bias. The way training data is chosen can accidentally cause bias. It can also originate from correlations: AI is used to classify individuals into groups and then make predictions based on the assumption that the individual will resemble the group’s members. This assumption may be incorrect in some instances. COMPAS, a commercial programme frequently used by US courts to assess the possibility of a defendant becoming a recidivist, is an example of this.

Despite the fact that the programme was not told the defendants’ races, ProPublica reports that the COMPAS-assigned recidivism risk rating of black defendants is considerably more likely to be an overestimate than that of white defendants. When AI is used for credit assessment or recruiting, for example, algorithmic bias might lead to biased results. Because of knowledge representation and knowledge engineering, AI machines can intelligently answer questions and make deductions about real-world facts. Ontology is a collection of objects, relations, concepts, and attributes that have been clearly established for software agents to understand. The most wide ontologies are upper ontologies, which attempt to offer a foundation for all other information and act as mediators between domain ontologies that cover specific knowledge about a given knowledge topic (field of interest or area of concern). The semantics of ontology are represented using description logic such as the web ontology language.