IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 9, September 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

आंतरिक सुरक्षा के लिए कृत्रिम बुद्धिमत्ता की उपयोगिता एवं नयी चुनौती

Dr. Amrendra Pratap Gond

Assistant Professor, Department of Defence and Strategic Studies, D.A.V. College, Kanpur,

Uttar Pradesh, India

सार

कृत्रिम बुद्धिमत्ता (एआई) मनुष्यों या जानवरों की बुद्धि के विपरीत मशीनों या सॉफ्टवेयर की बुद्धिमत्ता है। यह कंप्यूटर विज्ञान में अध्ययन का क्षेत्र भी है जो बुद्धिमान मशीनों का विकास और अध्ययन करता है। "एआई" स्वयं मशीनों को भी संदर्भित कर सकता है।

एआई तकनीक का व्यापक रूप से उद्योग, सरकार और विज्ञान में उपयोग किया जाता है। कुछ हाई-प्रोफ़ाइल एप्लिकेशन हैं: उन्नत वेब खोज इंजन (उदाहरण के लिए, Google खोज), अनुशंसा प्रणाली (यूट्यूब , अमेज़ॅन और नेटफ्लिक्स द्वारा उपयोग किया जाता है), मानव भाषण को समझना (जैसे सिरी और एलेक्सा), स्वयं-ड्राइविंग कारें (उदाहरण के लिए, वेमो)), उत्पादक या रचनात्मक उपकरण (चैटजीपीटी और एआई कला), और रणनीतिक खेलों में उच्चतम स्तर पर प्रतिस्पर्धा(जैसे शतरंज और गो)। [1]

परिचय

आर्टिफिशियल इंटेलिजेंस की स्थापना 1956 में एक अकादमिक अनुशासन के रूप में की गई थी। [2] यह क्षेत्र आशावाद के कई चक्रों से गुजरा [3] [4] जिसके बाद निराशा और धन की हानि हुई, [5] [6] लेकिन 2012 के बाद, जब गहन शिक्षा आगे निकल गई पिछली सभी एआई तकनीकों, [7] फंडिंग और रुचि में भारी वृद्धि हुई थी।

एआई अनुसंधान के विभिन्न उप-क्षेत्र विशेष लक्ष्यों और विशेष उपकरणों के उपयोग के आसपास केंद्रित हैं। एआई अनुसंधान के पारंपरिक लक्ष्यों में तर्क , ज्ञान प्रतिनिधित्व , योजना , सीखना , प्राकृतिक भाषा प्रसंस्करण , धारणा और रोबोटिक्स के लिए समर्थन शामिल हैं । [ए] सामान्य बुद्धिमत्ता (किसी मनमानी समस्या को हल करने की क्षमता) इस क्षेत्र के दीर्घकालिक लक्ष्यों में से एक है। [8] इन समस्याओं को हल करने के लिए, एआई शोधकर्ताओं ने खोज और गणितीय अनुकूलन, औपचारिक तर्क, कृत्रिम तंत्रिका नेटवर्क सहित समस्या-समाधान तकनीकों की एक विस्तृत श्रृंखला को अनुकूलित और एकीकृत किया है।, और सांख्यिकी , संभाव्यता और अर्थशास्त्र पर आधारित विधियाँ । [बी] एआई मनोविज्ञान , भाषा विज्ञान , दर्शन , तंत्रिका विज्ञान और कई अन्य क्षेत्रों पर भी आधारित है । [9]

लक्ष्य

बुद्धिमत्ता के अनुकरण (या निर्माण) की सामान्य समस्या को उप-समस्याओं में विभाजित किया गया है। इनमें विशेष लक्षण या क्षमताएं शामिल हैं जिन्हें शोधकर्ता एक बुद्धिमान प्रणाली से प्रदर्शित करने की अपेक्षा करते हैं। नीचे वर्णित लक्षणों पर सबसे अधिक ध्यान दिया गया है और ये एआई अनुसंधान के दायरे को कवर करते हैं। [ए]

तर्क, समस्या-समाधान

प्रारंभिक शोधकर्ताओं ने ऐसे एल्गोरिदम विकसित किए जो चरण-दर-चरण तर्क की नकल करते हैं जिसका उपयोग मनुष्य पहेलियाँ हल करते समय या तार्किक निष्कर्ष निकालते समय करते हैं। [10] 1980 और 1990 के दशक के अंत

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

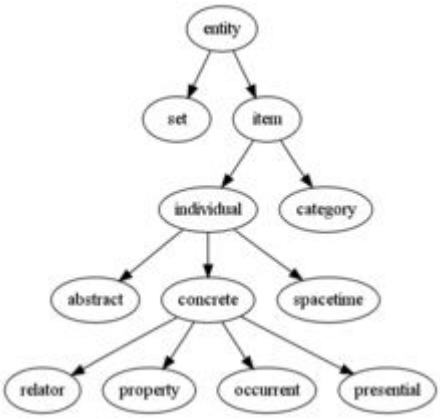
|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

तक, संभाव्यता और अर्थशास्त्र की अवधारणाओं को नियोजित करते हुए, अनिश्चित या अधूरी जानकारी से निपटने के तरीके विकसित किए गए थे । [11]

इनमें से कई एल्गोरिदम बड़ी तर्क समस्याओं को हल करने के लिए अपर्याप्त हैं क्योंकि वे एक "कॉम्बिनेटोरियल विस्फोट" का अनुभव करते हैं: जैसे-जैसे समस्याएं बड़ी होती गईं, वे तेजी से धीमी हो गईं। [12] यहां तक कि मनुष्य भी शायद ही कभी चरण-दर-चरण कटौती का उपयोग करते हैं जिसे प्रारंभिक एआई अनुसंधान मॉडल कर सकता है। वे तेज़, सहज निर्णयों का उपयोग करके अपनी अधिकांश समस्याओं का समाधान करते हैं। [13] सटीक और कुशल तर्क एक अनसुलझी समस्या है।

ज्ञान निरूपण



एक ऑन्कोलॉजी एक डोमेन के भीतर अवधारणाओं के एक सेट और उन अवधारणाओं के बीच संबंधों के रूप में ज्ञान का प्रतिनिधित्व करती है।

ज्ञान प्रतिनिधित्व और ज्ञान इंजीनियरिंग [14] एआई कार्यक्रमों को बुद्धिमानी से सवालों के जवाब देने और वास्तविक दुनिया के तथ्यों के बारे में निष्कर्ष निकालने की अनुमति देता है। औपचारिक ज्ञान अभ्यावेदन का उपयोग सामग्री-आधारित अनुक्रमण और पुनर्प्राप्ति, [15] दृश्य व्याख्या, [16] नैदानिक निर्णय समर्थन, [17] ज्ञान खोज (बड़े डेटाबेस से "दिलचस्प" और कार्रवाई योग्य निष्कर्ष निकालना), [18] और अन्य क्षेत्रों में किया जाता है। . [19]

ज्ञान का आधार ज्ञान का एक समूह है जिसे एक ऐसे रूप में दर्शाया जाता है जिसका उपयोग किसी प्रोग्राम द्वारा किया जा सकता है। ऑन्टोलॉजी ज्ञान के क्षेत्र द्वारा उपयोग की जाने वाली वस्तुओं, संबंधों, अवधारणाओं और गुणों का समूह है। [20] ज्ञान के आधारों को चीजों का प्रतिनिधित्व करने की आवश्यकता है जैसे: वस्तुएं, गुण, श्रेणियां और वस्तुओं के बीच संबंध; [21] स्थितियाँ, घटनाएँ, स्थितियाँ और समय; [22] कारण और प्रभाव; [23] ज्ञान के बारे में ज्ञान (हम जो जानते हैं उसके बारे में दूसरे लोग क्या जानते हैं); [24] डिफ़ॉल्ट तर्क (जिन चीज़ों को मनुष्य मान लेता है वे तब तक सत्य हैं जब तक कि उन्हें अलग तरह से नहीं बताया जाता है और तब भी सत्य रहेंगी जब अन्य तथ्य बदल रहे हों); [25]और ज्ञान के कई अन्य पहलू और क्षेत्र।

केआर में सबसे कठिन समस्याओं में से हैं: सामान्य ज्ञान ज्ञान की व्यापकता (परमाणु तथ्यों का समूह जो औसत व्यक्ति जानता है वह बहुत बड़ा है); [26] और अधिकांश सामान्य ज्ञान का उप-प्रतीकात्मक रूप (लोग जो कुछ भी जानते हैं उसे "तथ्यों" या "कथनों" के रूप में प्रस्तुत नहीं किया जाता है जिसे वे मौखिक रूप से व्यक्त कर सकते हैं)। [13]

एआई अनुप्रयोगों के लिए ज्ञान प्राप्त करना एक कठिन समस्या है। [सी] आधुनिक एआई इंटरनेट (विकिपीडिया सहित) को " स्क्रैप " करके ज्ञान इकट्ठा करता है। ज्ञान स्वयं स्वयंसेवकों और पेशेवरों द्वारा एकत्र किया गया था जिन्होंने जानकारी प्रकाशित की थी (जो एआई कंपनियों को अपना काम प्रदान करने के लिए सहमत हो भी सकते हैं और नहीं भी)। [29] यह " क्राउड सोर्स्ड " तकनीक यह गारंटी नहीं देती कि ज्ञान सही या विश्वसनीय है। बड़े भाषा मॉडल (जैसे चैट-जीपीटी) का ज्ञान अत्यधिक अविश्वसनीय है - यह गलत सूचना और झूठ उत्पन्न करता है (जिसे " मतिभ्रम" कहा जाता है)")। इन आधुनिक एआई अनुप्रयोगों के लिए सटीक ज्ञान प्रदान करना एक अनसुलझी समस्या है।

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

योजना बनाना और निर्णय लेना

एक "एजेंट" वह सब कुछ है जो दुनिया को समझता है और कार्रवाई करता है। एक तर्कसंगत एजेंट के पास लक्ष्य या प्राथमिकताएँ होती हैं और वह उन्हें पूरा करने के लिए कार्रवाई करता है। [डी] [30] स्वचालित योजना में , एजेंट का एक विशिष्ट लक्ष्य होता है। [31] स्वचालित निर्णय लेने में , एजेंट की प्राथमिकताएँ होती हैं - कुछ स्थितियाँ हैं जिनमें वह रहना पसंद करेगा, और कुछ स्थितियाँ हैं जिनसे वह बचने की कोशिश कर रहा है। निर्णय लेने वाला एजेंट प्रत्येक स्थिति के लिए एक नंबर निर्दिष्ट करता है (जिसे " उपयोगिता " कहा जाता है) जो मापता है कि एजेंट इसे कितना पसंद करता है। प्रत्येक संभावित क्रिया के लिए, यह " अपेक्षित उपयोगिता " की गणना कर सकता है: उपयोगिताकार्रवाई के सभी संभावित परिणामों को, परिणाम घटित होने की संभावना के आधार पर भारित किया जाता है। इसके बाद यह अधिकतम अपेक्षित उपयोगिता के साथ कार्रवाई चुन सकता है। [32]

शास्त्रीय नियोजन में , एजेंट को ठीक-ठीक पता होता है कि किसी भी कार्रवाई का प्रभाव क्या होगा। [33] हालाँकि, अधिकांश वास्तविक दुनिया की समस्याओं में, एजेंट उस स्थिति के बारे में निश्चित नहीं हो सकता है जिसमें वे हैं (यह "अज्ञात" या "अदृश्य" है) और यह निश्चित रूप से नहीं जान सकता है कि प्रत्येक संभावित कार्रवाई के बाद क्या होगा (यह "नियतात्मक" नहीं है)। इसे संभाव्य अनुमान लगाकर एक कार्रवाई का चयन करना चाहिए और फिर यह देखने के लिए स्थिति का पुनर्मूल्यांकन करना चाहिए कि क्या कार्रवाई काम कर रही है। [34] कुछ समस्याओं में, एजेंट की प्राथमिकताएं अनिश्चित हो सकती हैं, खासकर यदि इसमें अन्य एजेंट या इंसान शामिल हों। इन्हें सीखा जा सकता है (उदाहरण के लिए, उलटा सुदृढीकरण सीखने के साथ) या एजेंट अपनी प्राथमिकताओं में सुधार करने के लिए जानकारी मांग सकता है। [35] सूचना मूल्य सिद्धांत का उपयोग खोजपूर्ण या प्रायोगिक कार्यों के मूल्य को मापने के लिए किया जा सकता है। [36] संभावित भविष्य की कार्रवाइयों और स्थितियों का स्थान आम तौर पर काफी बड़ा होता है, इसलिए एजेंटों को कार्रवाई करनी चाहिए और परिस्थितियों का मूल्यांकन करना चाहिए, जबकि यह अनिश्चित होना चाहिए कि परिणाम क्या होगा।

मार्कोव निर्णय प्रक्रिया में एक संक्रमण मॉडल होता है जो इस संभावना का वर्णन करता है कि एक विशेष कार्रवाई राज्य को एक विशेष तरीके से बदल देगी, और एक इनाम फ़ंक्शन जो प्रत्येक राज्य की उपयोगिता और प्रत्येक कार्रवाई की लागत की आपूर्ति करता है। एक नीति प्रत्येक संभावित राज्य के साथ एक निर्णय को जोड़ती है। नीति की गणना की जा सकती है (उदाहरण के लिए पुनरावृत्ति द्वारा), अनुमानी हो सकती है , या इसे सीखा जा सकता है। [37]

गेम थ्योरी कई इंटरैक्टिंग एजेंटों के तर्कसंगत व्यवहार का वर्णन करती है, और इसका उपयोग एआई कार्यक्रमों में किया जाता है जो ऐसे निर्णय लेते हैं जिनमें अन्य एजेंट शामिल होते हैं। [38]

सीखना

मशीन लर्निंग उन प्रोग्रामों का अध्ययन है जो किसी दिए गए कार्य पर अपने प्रदर्शन को स्वचालित रूप से सुधार सकते हैं। [39] यह शुरू से ही एआई का हिस्सा रहा है। [इ]

मशीन लर्निंग कई प्रकार की होती है। बिना पर्यवेक्षित शिक्षण डेटा की एक धारा का विश्लेषण करता है और पैटर्न ढूंढता है और बिना किसी अन्य मार्गदर्शन के भविष्यवाणियां करता है। [42] पर्यवेक्षित शिक्षण के लिए मानव को पहले इनपुट डेटा को लेबल करने की आवश्यकता होती है, और यह दो मुख्य किस्मों में आता है: वर्गीकरण (जहां प्रोग्राम को यह अनुमान लगाना सीखना चाहिए कि इनपुट किस श्रेणी में है) और रिग्रेशन (जहां प्रोग्राम को एक संख्यात्मक फ़ंक्शन के आधार पर अनुमान लगाना होगा) संख्यात्मक इनपुट पर)। [43] सुदृढीकरण सीखने में एजेंट को अच्छी प्रतिक्रियाओं के लिए पुरस्कृत किया जाता है और खराब प्रतिक्रियाओं के लिए दंडित किया जाता है। एजेंट उन प्रतिक्रियाओं को चुनना सीखता है जिन्हें "अच्छे" के रूप में वर्गीकृत किया गया है। [44] स्थानांतरण शिक्षणजब एक समस्या से प्राप्त ज्ञान को एक नई समस्या पर लागू किया जाता है। [45] गहन शिक्षण इन सभी प्रकार के शिक्षण के लिए कृत्रिम तंत्रिका नेटवर्क का उपयोग करता है ।

कम्प्यूटेशनल शिक्षण सिद्धांत कम्प्यूटेशनल जटिलता , नमूना जटिलता (कितना डेटा आवश्यक है), या अनुकूलन की अन्य धारणाओं द्वारा शिक्षार्थियों का आकलन कर सकता है । [46]

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

प्राकृतिक भाषा प्रसंस्करण

प्राकृतिक भाषा प्रसंस्करण (एनएलपी) [47] प्रोग्रामों को अंग्रेजी जैसी मानव भाषाओं में पढ़ने, लिखने और संचार करने की अनुमति देता है। विशिष्ट समस्याओं में वाक् पहचान , वाक् संश्लेषण , मशीनी अनुवाद , सूचना निष्कर्षण , सूचना पुनर्प्राप्ति और प्रश्न उत्तर शामिल हैं । [48]

नोम चॉम्स्की के जेनरेटिव व्याकरण और सिमेंटिक नेटवर्क पर आधारित प्रारंभिक कार्य में शब्द-बोध की व्याख्या में कठिनाई होती थी [एफ] जब तक कि यह " सूक्ष्म-दुनिया " नामक छोटे डोमेन तक सीमित न हो (सामान्य ज्ञान ज्ञान समस्या के कारण [26])।

एनएलपी के लिए आधुनिक गहन शिक्षण तकनीकों में शब्द एम्बेडिंग (कितनी बार एक शब्द दूसरे के करीब आता है), [49] ट्रांसफार्मर (जो पाठ में पैटर्न ढूंढता है), [50] और अन्य शामिल हैं। [51] 2019 में, जेनरेटिव पूर्व-प्रशिक्षित ट्रांसफार्मर (या "जीपीटी") भाषा मॉडल ने सुसंगत पाठ उत्पन्न करना शुरू किया, [52] [53] और 2023 तक ये मॉडल बार परीक्षा , एसएटी पर मानव-स्तर के स्कोर प्राप्त करने में सक्षम थे। , जीआरई , और कई अन्य वास्तविक दुनिया के अनुप्रयोग। [54]

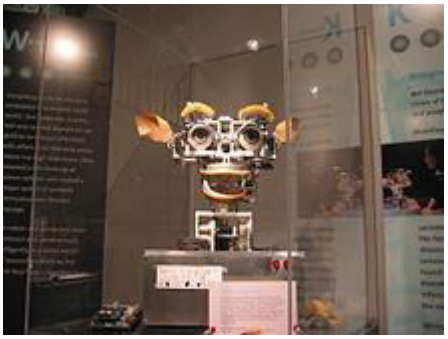
धारणा

फ़ीचर डिटेक्शन (चित्रित: एज डिटेक्शन) एआई को कच्चे डेटा से सूचनात्मक सार संरचनाएं बनाने में मदद करता है।

मशीन धारणा दुनिया के पहलुओं का पता लगाने के लिए सेंसर (जैसे कैमरा, माइक्रोफोन, वायरलेस सिग्नल, सक्रिय लिडार , सोनार, रडार और स्पर्श सेंसर) से इनपुट का उपयोग करने की क्षमता है। कंप्यूटर विज़न दृश्य इनपुट का विश्लेषण करने की क्षमता है। [55] इस क्षेत्र में वाक् पहचान , [56] छवि वर्गीकरण , [57] चेहरे की पहचान , वस्तु पहचान , [58] और रोबोटिक धारणा शामिल है । [59]

विचार-विमर्श

सामाजिक बुद्धिमत्ता



किस्मत , प्रारंभिक सामाजिक कौशल वाला एक रोबोट [61]

प्रभावशाली कंप्यूटिंग एक अंतःविषय छत्र है जिसमें ऐसी प्रणालियाँ शामिल हैं जो मानवीय भावना, भावना और मनोदशा को पहचानती हैं, व्याख्या करती हैं, संसाधित करती हैं या अनुकरण करती हैं । [62] उदाहरण के लिए, कुछ आभासी सहायकों को बातचीत करने या यहां तक कि विनोदी ढंग से मजाक करने के लिए प्रोग्राम किया जाता है; यह उन्हें मानव संपर्क की

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

भावनात्मक गतिशीलता के प्रति अधिक संवेदनशील बनाता है, या अन्यथा मानव-कंप्यूटर संपर्क को सुविधाजनक बनाता है । हालाँकि, इससे भोले-भाले उपयोगकर्ताओं को यह अवास्तविक अवधारणा मिलती है कि मौजूदा कंप्यूटर एजेंट वास्तव में कितने बुद्धिमान हैं। [63] भावात्मक कंप्यूटिंग से संबंधित मध्यम सफलताओं में पाठ्य भावना विश्लेषण और, हाल ही में, शामिल हैं।मल्टीमॉडल भावना विश्लेषण , जिसमें एआई वीडियोटेप किए गए विषय द्वारा प्रदर्शित प्रभावों को वर्गीकृत करता है। [64]

सामान्य बुद्धि

कृत्रिम सामान्य बुद्धि वाली एक मशीन मानव बुद्धि के समान व्यापकता और बहुमुखी प्रतिभा के साथ विभिन्न प्रकार की समस्याओं को हल करने में सक्षम होनी चाहिए। [8]

औजार

एआई अनुसंधान उपरोक्त लक्ष्यों को पूरा करने के लिए विभिन्न प्रकार के उपकरणों का उपयोग करता है। [बी]

खोज और अनुकूलन

एआई कई संभावित समाधानों को समझदारी से खोजकर कई समस्याओं का समाधान कर सकता है। [65] एआई में दो बहुत अलग प्रकार की खोज का उपयोग किया जाता है: राज्य अंतरिक्ष खोज और स्थानीय खोज ।

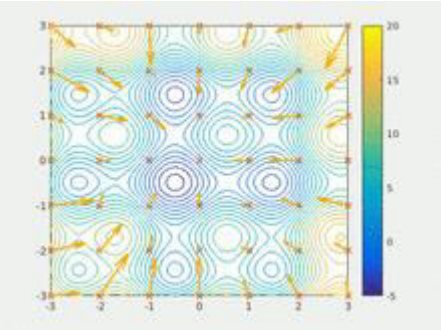
राज्य अंतरिक्ष खोज

राज्य अंतरिक्ष खोज एक लक्ष्य राज्य को खोजने का प्रयास करने के लिए संभावित राज्यों के पेड़ के माध्यम से खोज करती है। [66] उदाहरण के लिए, योजना एल्गोरिदम लक्ष्यों और उप-लक्ष्यों के पेड़ों के माध्यम से खोज करते हैं, लक्ष्य लक्ष्य के लिए रास्ता खोजने का प्रयास करते हैं, इस प्रक्रिया को साधन-अंत विश्लेषण कहा जाता है । [67]

वास्तविक दुनिया की अधिकांश समस्याओं के लिए सरल विस्तृत खोजें [68] शायद ही कभी पर्याप्त होती हैं: खोज स्थान (खोज करने के लिए स्थानों की संख्या) तेजी से खगोलीय संख्या तक बढ़ जाती है । परिणाम एक ऐसी खोज है जो बहुत धीमी है या कभी पूरी नहीं होती। [12] " ह्युरिस्टिक्स " या "अंगूठे के नियम" उन विकल्पों को प्राथमिकता देने में मदद कर सकते हैं जो किसी लक्ष्य तक पहुंचने की अधिक संभावना रखते हैं। [69]

शतरंज या गो जैसे खेल-खेलने वाले कार्यक्रमों के लिए प्रतिकूल खोज का उपयोग किया जाता है। यह संभावित चालों और जवाबी चालों के वृक्ष के माध्यम से विजयी स्थिति की तलाश करता है। [70]

स्थानीय खोज



वैश्विक न्यूनतम की तलाश में एक कण झुंड

स्थानीय खोज किसी समस्या का संख्यात्मक समाधान खोजने के लिए गणितीय अनुकूलन का उपयोग करती है । यह किसी प्रकार के अनुमान से शुरू होता है और फिर अनुमान को धीरे-धीरे परिष्कृत करता जाता है जब तक कि कोई और सुधार न किया जा सके। इन एल्गोरिदम को अंधी पहाड़ी चढ़ाई के रूप में देखा जा सकता है : हम परिदृश्य पर एक यादृच्छिक बिंदु पर

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

खोज शुरू करते हैं, और फिर, छलांग या कदमों से, हम अपने अनुमान को ऊपर की ओर बढ़ाते रहते हैं, जब तक कि हम शीर्ष पर नहीं पहुंच जाते। इस प्रक्रिया को स्टोकेस्टिक ग्रेडिएंट डिसेंट कहा जाता है । [71]

विकासवादी संगणना अनुकूलन खोज के एक रूप का उपयोग करती है। उदाहरण के लिए, वे जीवों की आबादी (अनुमानों) से शुरू कर सकते हैं और फिर उन्हें उत्परिवर्तित और पुनर्संयोजित करने की अनुमति दे सकते हैं, प्रत्येक पीढ़ी में जीवित रहने के लिए केवल सबसे योग्य का चयन कर सकते हैं (अनुमानों को परिष्कृत करते हुए)। [72]

वितरित खोज प्रक्रियाएं झुंड खुफिया एल्गोरिदम के माध्यम से समन्वयित कर सकती हैं। खोज में उपयोग किए जाने वाले दो लोकप्रिय झुंड एल्गोरिदम कण झुंड अनुकूलन (पक्षियों के झुंड से प्रेरित) और चींटी कॉलोनी अनुकूलन (चींटी ट्रेल्स से प्रेरित) हैं। [73]

तंत्रिका नेटवर्क और सांख्यिकीय वर्गीकरणकर्ता (नीचे चर्चा की गई है), स्थानीय खोज के एक रूप का भी उपयोग करते हैं, जहां खोजा जाने वाला "परिदृश्य" सीखने से बनता है।

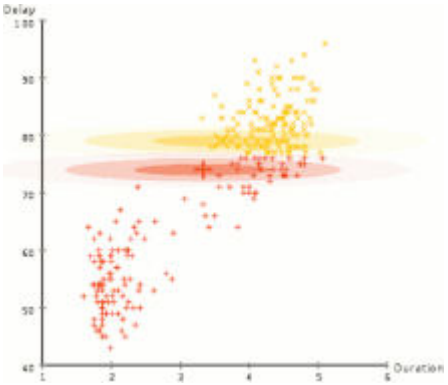
तर्क

औपचारिक तर्क का उपयोग तर्क और ज्ञान प्रतिनिधित्व के लिए किया जाता है । [74] औपचारिक तर्क दो मुख्य रूपों में आता है: प्रस्तावात्मक तर्क (जो सही या गलत बयानों पर काम करता है और "और", "या", "नहीं" और "निहित" जैसे तार्किक संयोजकों का उपयोग करता है) [75] और विधेय तर्क (जो वस्तुओं, विधेय और संबंधों पर भी काम करता है और " हर एक्स एक वाई है " और " कुछ एक्स हैं जो वाई हैं " जैसे क्वांटिफायर का उपयोग करता है)। [76]

तार्किक अनुमान (या कटौती) अन्य कथनों से एक नए कथन (निष्कर्ष) को साबित करने की प्रक्रिया है जो पहले से ही सत्य (परिसर) के रूप में जाना जाता है। [77] एक तार्किक ज्ञान आधार अनुमान के विशेष मामले के रूप में प्रश्नों और दावों को भी संभालता है। [78] एक अनुमान नियम बताता है कि प्रमाण में एक वैध कदम क्या है। सबसे सामान्य अनुमान नियम संकल्प है । [79] अनुमान को उस पथ को खोजने के लिए खोज करने तक सीमित किया जा सकता है जो परिसर से निष्कर्ष तक जाता है, जहां प्रत्येक चरण एक का अनुप्रयोग हैअनुमान नियम . [80] प्रतिबंधित डोमेन में संक्षिप्त प्रमाणों को छोड़कर इस तरह से किया गया अनुमान कठिन है। कोई कारगर, सशक्त एवं सामान्य विधि नहीं खोजी जा सकी है। [81]

फ़ज़ी लॉजिक 0 और 1 के बीच "सच्चाई की डिग्री" निर्दिष्ट करता है और अनिश्चितता और संभाव्य स्थितियों को संभालता है। [82] गैर-मोनोटोनिक तर्क डिफ़ॉल्ट तर्क को संभालने के लिए डिज़ाइन किए गए हैं । [25] कई जटिल डोमेन का वर्णन करने के लिए तर्क के अन्य विशेष संस्करण विकसित किए गए हैं (ऊपर ज्ञान प्रतिनिधित्व देखें)।

अनिश्चित तर्क के लिए संभाव्य तरीके



ओल्ड फेथफुल विस्फोट डेटा की उम्मीद-अधिकतम क्लस्टरिंग एक यादृच्छिक अनुमान से शुरू होती है लेकिन फिर विस्फोट के दो भौतिक रूप से अलग-अलग तरीकों के सटीक क्लस्टरिंग पर सफलतापूर्वक एकत्रित होती है।

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

एआई में कई समस्याओं (तर्क, योजना, सीखने, धारणा और रोबोटिक्स सहित) के लिए एजेंट को अधूरी या अनिश्चित जानकारी के साथ काम करने की आवश्यकता होती है। एआई शोधकर्ताओं ने संभाव्यता सिद्धांत और अर्थशास्त्र के तरीकों का उपयोग करके इन समस्याओं को हल करने के लिए कई उपकरण तैयार किए हैं । [83]

बायेसियन नेटवर्क [84] एक बहुत ही सामान्य उपकरण है जिसका उपयोग कई समस्याओं के लिए किया जा सकता है, जिसमें तर्क (बायेसियन अनुमान एल्गोरिदम का उपयोग करना), [जी] [86] सीखना (अपेक्षा-अधिकतमकरण एल्गोरिदम का उपयोग करना), [एच] [88] शामिल है । योजना (निर्णय नेटवर्क का उपयोग करके) [89] और धारणा (गतिशील बायेसियन नेटवर्क का उपयोग करके)। [90]

संभाव्य एल्गोरिदम का उपयोग डेटा की धाराओं को फ़िल्टर करने, भविष्यवाणी करने, सुचारू करने और स्पष्टीकरण खोजने के लिए भी किया जा सकता है, जिससे धारणा प्रणालियों को समय के साथ होने वाली प्रक्रियाओं का विश्लेषण करने में मदद मिलती है (उदाहरण के लिए, छिपे हुए मार्कोव मॉडल या कलमन फ़िल्टर)। [90]

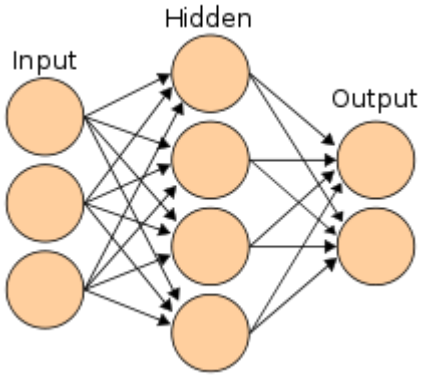
सटीक गणितीय उपकरण विकसित किए गए हैं जो निर्णय सिद्धांत , निर्णय विश्लेषण , [91] और सूचना मूल्य सिद्धांत का उपयोग करके विश्लेषण करते हैं कि एक एजेंट कैसे विकल्प चुन सकता है और योजना बना सकता है । [92] इन उपकरणों में मार्कोव निर्णय प्रक्रियाएं , [93] गतिशील निर्णय नेटवर्क , [90] गेम सिद्धांत और तंत्र डिजाइन जैसे मॉडल शामिल हैं । [94]

वर्गीकरणकर्ता और सांख्यिकीय सीखने के तरीके

सबसे सरल एआई अनुप्रयोगों को दो प्रकारों में विभाजित किया जा सकता है: एक तरफ क्लासिफायर (उदाहरण के लिए "यदि चमकदार है तो हीरा"), और दूसरी ओर नियंत्रक (उदाहरण के लिए "यदि हीरा है तो हीरा उठाएं")। क्लासिफायर [95] ऐसे फ़ंक्शन हैं जो निकटतम मिलान निर्धारित करने के लिए पैटर्न मिलान का उपयोग करते हैं। पर्यवेक्षित शिक्षण का उपयोग करके चुने हुए उदाहरणों के आधार पर उन्हें ठीक किया जा सकता है । प्रत्येक पैटर्न (जिसे " अवलोकन " भी कहा जाता है) को एक निश्चित पूर्वनिर्धारित वर्ग के साथ लेबल किया जाता है। सभी अवलोकनों को उनके वर्ग लेबल के साथ मिलाकर डेटा सेट के रूप में जाना जाता है । जब कोई नया अवलोकन प्राप्त होता है, तो उस अवलोकन को पिछले अनुभव के आधार पर वर्गीकृत किया जाता है। [43]

कई प्रकार के क्लासिफायर उपयोग में हैं। निर्णय वृक्ष सबसे सरल और सबसे व्यापक रूप से उपयोग किया जाने वाला प्रतीकात्मक मशीन लर्निंग एल्गोरिदम है। [96] 1990 के दशक के मध्य तक के-निकटतम पड़ोसी एल्गोरिदम सबसे व्यापक रूप से इस्तेमाल किया जाने वाला एनालॉग एआई था, और सपोर्ट वेक्टर मशीन (एसवीएम) जैसी कर्नेल विधियों ने 1990 के दशक में के-निकटतम पड़ोसी को विस्थापित कर दिया। [97] कथित तौर पर अनुभवहीन बेयस क्लासिफायरियर अपनी स्केलेबिलिटी के कारण Google में "सबसे व्यापक रूप से उपयोग किया जाने वाला शिक्षार्थी" है [98] । [99] तंत्रिका नेटवर्क का उपयोग क्लासिफायर के रूप में भी किया जाता है। [100]

कृत्रिम तंत्रिका प्रसार



तंत्रिका नेटवर्क नोड्स का एक परस्पर जुड़ा हुआ समूह है, जो मानव मस्तिष्क में न्यूरॉन्स के विशाल नेटवर्क के समान है ।

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

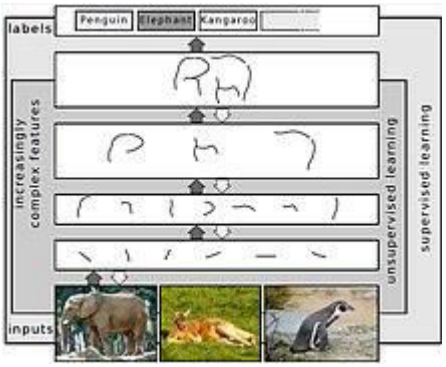
| DOI: 10.15680/IJIRCCE.2023.1109018 |

कृत्रिम तंत्रिका नेटवर्क [100] मानव मस्तिष्क के डिजाइन से प्रेरित थे: एक सरल "न्यूरॉन" एन अन्य न्यूरॉन्स से इनपुट स्वीकार करता है, जिनमें से प्रत्येक, सक्रिय होने (या "निकालने") के पक्ष या विपक्ष में एक महत्वपूर्ण "वोट" डालता है। क्या न्यूरॉन एन को स्वयं सक्रिय होना चाहिए। व्यवहार में, इनपुट "न्यूरॉन्स" संख्याओं की एक सूची है , "वजन" एक मैट्रिक्स है , अगली परत एक बढ़ते फ़ंक्शन द्वारा स्केल किया गया डॉट उत्पाद (यानी, कई भारित योग) है , जैसे कि लॉजिस्टिक फ़ंक्शन । "वास्तविक तंत्रिका कोशिकाओं और संरचनाओं से समानता सतही है",. [101] [मैं]

तंत्रिका नेटवर्क के लिए सीखने के एल्गोरिदम स्थानीय खोज का उपयोग करके उन वजनों को चुनते हैं जो प्रशिक्षण के दौरान प्रत्येक इनपुट के लिए सही आउटपुट प्राप्त करेंगे। सबसे आम प्रशिक्षण तकनीक बैकप्रॉपैगेशन एल्गोरिदम है। [102] तंत्रिका नेटवर्क इनपुट और आउटपुट के बीच जटिल संबंधों को मॉडल करना और डेटा में पैटर्न ढूंढना सीखते हैं। सिद्धांत रूप में, एक तंत्रिका नेटवर्क कोई भी कार्य सीख सकता है। [103]

फीडफॉरवर्ड न्यूरल नेटवर्क में सिग्नल केवल एक ही दिशा में गुजरता है। [104] आवर्तक तंत्रिका नेटवर्क आउटपुट सिग्नल को वापस इनपुट में फ़ीड करते हैं, जो पिछले इनपुट घटनाओं की अल्पकालिक यादों की अनुमति देता है। दीर्घकालिक अल्पकालिक मेमोरी आवर्ती नेटवर्क के लिए सबसे सफल नेटवर्क आर्किटेक्चर है। [105] परसेप्ट्रोन [106] न्यूरॉन्स की केवल एक परत का उपयोग करते हैं, गहन शिक्षा [107] कई परतों का उपयोग करती है। संवादात्मक तंत्रिका नेटवर्क उन न्यूरॉन्स के बीच संबंध को मजबूत करते हैं जो एक दूसरे के "करीब" होते हैं - यह छवि प्रसंस्करण में विशेष रूप से महत्वपूर्ण है , जहां न्यूरॉन्स का एक स्थानीय सेट होना चाहिएनेटवर्क द्वारा किसी वस्तु की पहचान करने से पहले एक "किनारे" की पहचान करें । [108]

ध्यान लगा के पढ़ना या सीखना



गहन शिक्षण में अमूर्तता की कई परतों पर छवियों का प्रतिनिधित्व करना [109]

डीप लर्निंग [107] नेटवर्क के इनपुट और आउटपुट के बीच न्यूरॉन्स की कई परतों का उपयोग करता है। कई परतें कच्चे इनपुट से उत्तरोत्तर उच्च-स्तरीय सुविधाएँ निकाल सकती हैं। उदाहरण के लिए, छवि प्रसंस्करण में , निचली परतें किनारों की पहचान कर सकती हैं, जबकि उच्च परतें मानव से संबंधित अवधारणाओं जैसे अंक या अक्षर या चेहरे की पहचान कर सकती हैं। [110]

डीप लर्निंग ने कृत्रिम बुद्धिमत्ता के कई महत्वपूर्ण उपक्षेत्रों में कार्यक्रमों के प्रदर्शन में काफी सुधार किया है, जिसमें कंप्यूटर विज़न , वाक् पहचान , छवि वर्गीकरण [111] और अन्य शामिल हैं। इतने सारे अनुप्रयोगों में गहन शिक्षण इतना अच्छा प्रदर्शन करता है इसका कारण 2023 तक ज्ञात नहीं है। [112] 2012-2015 में गहन शिक्षण की अचानक सफलता किसी नई खोज या सैद्धांतिक सफलता (गहरे तंत्रिका नेटवर्क और बैकप्रोपेगेशन) के कारण नहीं हुई थी 1950 के दशक तक कई लोगों द्वारा वर्णित किया गया है) [जे] लेकिन दो कारकों के कारण: कंप्यूटर शक्ति में अविश्वसनीय वृद्धि (जीपीयू पर स्विच करने से गति में सौ गुना वृद्धि सहित)) और बड़ी मात्रा में प्रशिक्षण डेटा की उपलब्धता, विशेष रूप से बेंचमार्क परीक्षण के लिए उपयोग किए जाने वाले विशाल क्यूरेटेड डेटासेट , जैसे कि इमेजनेट । [क]

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

विशिष्ट हार्डवेयर और सॉफ्टवेयर

2010 के अंत में, ग्राफिक्स प्रोसेसिंग यूनिट (जीपीयू) जो तेजी से एआई-विशिष्ट संवर्द्धन के साथ डिजाइन किए गए थे और विशेष टेन्सरफ्लो सॉफ्टवेयर के साथ उपयोग किए गए थे, ने पहले इस्तेमाल किए गए सेंट्रल प्रोसेसिंग यूनिट (सीपीयू) को बड़े पैमाने पर (वाणिज्यिक और शैक्षणिक) के लिए प्रमुख साधन के रूप में बदल दिया था। मशीन लर्निंग मॉडल का प्रशिक्षण। [121] ऐतिहासिक रूप से, लिस्प , प्रोलॉग और अन्य जैसी विशिष्ट भाषाओं का उपयोग किया गया था।

अनुप्रयोग

कलाकार जोसेफ आयरले के इस 2018 प्रोजेक्ट के लिए एआई को पुनर्जागरण चित्रकार राफेल के रंगों और ब्रशस्ट्रोक में विशिष्ट पैटर्न सीखना था । चित्र में अभिनेत्री ओरनेला मुटी का चेहरा दिखाया गया है , जिसे राफेल की शैली में एआई द्वारा "चित्रित" किया गया है।

एआई और मशीन लर्निंग तकनीक का उपयोग 2020 के अधिकांश आवश्यक अनुप्रयोगों में किया जाता है, जिनमें शामिल हैं: खोज इंजन (जैसे Google खोज), ऑनलाइन विज्ञापनों को लक्षित करना , [122] अनुशंसा प्रणाली (नेटफ्लिक्स , यूट्यूब या अमेज़ॅन द्वारा प्रस्तावित), इंटरनेट ट्रैफ़िक चलाना , [123] [124] लक्षित विज्ञापन (एडसेंस , फेसबुक), आभासी सहायक (जैसे सिरी या एलेक्सा), [125] स्वायत्त वाहन (सहित)ड्रोन , एडीएएस और सेल्फ-ड्राइविंग कारें), स्वचालित भाषा अनुवाद (माइक्रोसॉफ्ट ट्रांसलेटर , गूगल ट्रांसलेट), चेहरे की पहचान (एप्पल का फेस आईडी या माइक्रोसॉफ्ट का डीपफेस) और छवि लेबलिंग (फेसबुक , एप्पल के आईफोटो और टिकटॉक द्वारा उपयोग किया जाता है) .

ऐसे हजारों सफल AI अनुप्रयोग भी हैं जिनका उपयोग विशिष्ट उद्योगों या संस्थानों की विशिष्ट समस्याओं को हल करने के लिए किया जाता है। 2017 के सर्वेक्षण में, पांच में से एक कंपनी ने बताया कि उन्होंने कुछ पेशकशों या प्रक्रियाओं में "एआई" को शामिल किया है। [126] कुछ उदाहरण ऊर्जा भंडारण , [127] चिकित्सा निदान, सैन्य रसद, न्यायिक निर्णयों के परिणाम की भविष्यवाणी करने वाले अनुप्रयोग, [128] विदेश नीति , [129] या आपूर्ति श्रृंखला प्रबंधन हैं।

एआई की सबसे उन्नत तकनीकों को प्रदर्शित करने और परीक्षण करने के लिए 1950 के दशक से गेम खेलने के कार्यक्रमों का उपयोग किया जाता रहा है। डीप ब्लू 11 मई 1997 को मौजूदा विश्व शतरंज चैंपियन, गैरी कास्परोव को हराने वाला पहला कंप्यूटर शतरंज-खेल प्रणाली बन गया। [130] 2011 में, एक ख़तरे में! क्विज़ शो प्रदर्शनी मैच, आईबीएम के प्रश्न उत्तर प्रणाली , वॉटसन , ने दो सबसे बड़े ख़तरे को हराया! चैंपियन, ब्रैड रटर और केन जेनिंग्स , महत्वपूर्ण अंतर से। [131] मार्च 2016 में, अल्फ़ागो ने गो के 5 में से 4 गेम जीतेगो चैंपियन ली सेडोल के साथ एक मैच में , बिना किसी बाधा के पेशेवर गो खिलाड़ी को हराने वाला पहला कंप्यूटर गो -प्लेइंग सिस्टम बन गया । [132] फिर इसने 2017 में के जी को हराया , जो उस समय लगातार दो वर्षों तक विश्व नंबर 1 रैंकिंग पर बने रहे। [133] [134] [135] अन्य कार्यक्रम अपूर्ण जानकारी वाले खेलों को संभालते हैं; जैसे कि अलौकिक स्तर पर पोकर के लिए, प्लुरिबस [एल] और सेफियस । [137] 2010 के दशक में डीपमाइंड ने एक "सामान्यीकृत कृत्रिम बुद्धिमत्ता" विकसित की जो कई विविध अटारी सीख सकती थी खेल अपने आप। [138]

2020 की शुरुआत में, जेनरेटिव एआई को व्यापक प्रसिद्धि मिली। GPT-3 और अन्य बड़े भाषा मॉडल पर आधारित ChatGPT को 14% अमेरिकी वयस्कों द्वारा आज़माया गया था। [139] मिडजर्नी , डीएएलएल -ई और स्टेबल डिफ्यूज़न जैसे एआई-आधारित टेक्स्ट-टू-इमेज जनरेटर के बढ़ते यथार्थवाद और उपयोग में आसानी ने वायरल एआई-जनरेटेड तस्वीरों का चलन बढ़ा दिया है । सफ़ेद पफ़र कोट पहने पोप फ्रांसिस की एक नकली तस्वीर ने व्यापक ध्यान आकर्षित किया , [142] डोनाल्ड ट्रम्प की काल्पनिक गिरफ्तारी ,[143] और पेंटागन पर हमले की अफवाह, [144] साथ ही पेशेवर रचनात्मक कलाओं में उपयोग। [145] [146]

अल्फाफोल्ड 2 (2020) ने प्रोटीन की 3डी संरचना को महीनों के बजाय घंटों में अनुमानित करने की क्षमता का प्रदर्शन किया। [147]

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

नीति

किसी भी शक्तिशाली तकनीक की तरह AI के भी संभावित लाभ और संभावित जोखिम हैं। एआई विज्ञान को आगे बढ़ाने और गंभीर समस्याओं का समाधान खोजने में सक्षम हो सकता है: डीप माइंड के डेमिस हसाबिस को उम्मीद है कि "बुद्धिमत्ता को हल किया जाएगा, और फिर उसका उपयोग बाकी सभी चीजों को हल करने के लिए किया जाएगा"। [148] हालाँकि, जैसे-जैसे एआई का उपयोग व्यापक हो गया है, कई अनपेक्षित परिणामों और जोखिमों की पहचान की गई है। [149]

परिणाम

यदि मशीन लर्निंग एप्लिकेशन पक्षपाती डेटा से सीखते हैं तो वे पक्षपातपूर्ण होंगे। [150] डेवलपर्स को इस बात की जानकारी नहीं होगी कि पूर्वाग्रह मौजूद है। [151] जिस तरह से प्रशिक्षण डेटा का चयन किया जाता है और जिस तरह से एक मॉडल को तैनात किया जाता है, उससे पूर्वाग्रह का परिचय दिया जा सकता है। [152] [150] यदि पक्षपातपूर्ण एल्गोरिदम का उपयोग ऐसे निर्णय लेने के लिए किया जाता है जो लोगों को गंभीर रूप से नुकसान पहुंचा सकता है (जैसा कि यह चिकित्सा , वित्त , भर्ती , आवास या पुलिसिंग में हो सकता है) तो एल्गोरिदम भेदभाव का कारण बन सकता है । [153] निष्पक्षतामशीन लर्निंग में इस बात का अध्ययन किया जाता है कि एल्गोरिथम पूर्वाग्रह से होने वाले नुकसान को कैसे रोका जाए। यह एआई के भीतर अकादमिक अध्ययन का गंभीर क्षेत्र बन गया है। शोधकर्ताओं ने पता लगाया है कि "निष्पक्षता" को सभी हितधारकों को संतुष्ट करने वाले तरीके से परिभाषित करना हमेशा संभव नहीं होता है। [154]

28 जून 2015 को, Google फ़ोटो की नई छवि लेबलिंग सुविधा ने गलती से जैकी एल्सीन और एक दोस्त को "गोरिल्ला" के रूप में पहचान लिया क्योंकि वे काले थे। सिस्टम को एक डेटासेट पर प्रशिक्षित किया गया था जिसमें काले लोगों की बहुत कम छवियां थीं, [155] एक समस्या जिसे "नमूना आकार असमानता" कहा जाता था। [156] Google ने सिस्टम को किसी भी चीज़ को "गोरिल्ला" के रूप में लेबल करने से रोककर इस समस्या को "ठीक" कर दिया। आठ साल बाद, 2023 में, Google फ़ोटो अभी भी गोरिल्ला की पहचान नहीं कर सका, और न ही Apple, Facebook, Microsoft और Amazon के समान उत्पाद। [157]

कम्पास एक व्यावसायिक कार्यक्रम है जिसका व्यापक रूप से अमेरिकी अदालतों द्वारा प्रतिवादी के दोबारा अपराध करने वाले बनने की संभावना का आकलन करने के लिए उपयोग किया जाता है । 2016 में, प्रोपब्लिका में जूलिया एंगविन ने पाया कि कम्पास ने नस्लीय पूर्वाग्रह प्रदर्शित किया, इस तथ्य के बावजूद कि कार्यक्रम में प्रतिवादियों की नस्ल नहीं बताई गई थी। हालाँकि श्वेत और अश्वेत दोनों के लिए त्रुटि दर बिल्कुल 61% के बराबर अंशांकित की गई थी, प्रत्येक जाति के लिए त्रुटियाँ अलग-अलग थीं - सिस्टम ने लगातार इस संभावना को कम करके आंका कि एक काला व्यक्ति फिर से अपराध करेगा और इस संभावना को कम करके आंका जाएगा कि एक श्वेत व्यक्ति ऐसा करेगा। दोबारा अपमान न करें. [158] 2017 में, कई शोधकर्ता [एम]दिखाया गया कि कम्पास के लिए निष्पक्षता के सभी संभावित उपायों को समायोजित करना गणितीय रूप से असंभव था जब डेटा में गोरों और अश्वेतों के लिए दोबारा अपराध की आधार दरें अलग-अलग थीं। [160]

एक प्रोग्राम पक्षपातपूर्ण निर्णय ले सकता है, भले ही डेटा स्पष्ट रूप से किसी समस्याग्रस्त विशेषता (जैसे "जाति" या "लिंग") का उल्लेख न करता हो। यह सुविधा अन्य सुविधाओं (जैसे "पता", "खरीदारी इतिहास" या "प्रथम नाम") के साथ सहसंबद्ध होगी, और कार्यक्रम इन सुविधाओं के आधार पर वही निर्णय लेगा जो वह "जाति" या "लिंग" पर करेगा। [161] मोरित्ज़ हार्डट ने कहा, "इस शोध क्षेत्र में सबसे मजबूत तथ्य यह है कि अंधेपन के माध्यम से निष्पक्षता काम नहीं करती है।" [162]

कम्पास की आलोचना ने एआई के दुरुपयोग की एक गहरी समस्या पर प्रकाश डाला। मशीन लर्निंग मॉडल "भविष्यवाणियां" करने के लिए डिज़ाइन किए गए हैं जो केवल तभी मान्य होते हैं जब हम मानते हैं कि भविष्य अतीत जैसा होगा। यदि उन्हें उस डेटा पर प्रशिक्षित किया जाता है जिसमें अतीत में नस्लवादी निर्णयों के परिणाम शामिल हैं, तो मशीन लर्निंग मॉडल को यह अनुमान लगाना होगा कि भविष्य में नस्लवादी निर्णय लिए जाएंगे। दुर्भाग्य से, यदि कोई एप्लिकेशन इन भविष्यवाणियों को अनुशंसाओं के रूप में उपयोग करता है , तो इनमें से कुछ "सिफारिशें" संभवतः नस्लवादी होंगी। [163] इस प्रकार, मशीन लर्निंग उन क्षेत्रों में निर्णय लेने में मदद करने के लिए उपयुक्त नहीं है जहां उम्मीद है कि भविष्य अतीत से बेहतर होगा। यह आवश्यक रूप से वर्णनात्मक है न कि निषेधात्मक। [एन]

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

पूर्वाग्रह और अनुचितता का पता नहीं चल पाता क्योंकि डेवलपर्स में अधिकतर श्वेत और पुरुष हैं: एआई इंजीनियरों में, लगभग 4% अश्वेत हैं और 20% महिलाएं हैं। [156]

दक्षिण कोरिया के सियोल में एसोसिएशन फॉर कंप्यूटिंग मशीनरी ने निष्पक्षता, जवाबदेही और पारदर्शिता (ACM FAccT 2022) पर अपने 2022 सम्मेलन में यह सिफारिश करते हुए निष्कर्ष प्रस्तुत और प्रकाशित किए कि जब तक AI और रोबोटिक्स सिस्टम पूर्वाग्रह गलतियों से मुक्त साबित नहीं हो जाते, तब तक वे असुरक्षित और त्रुटिपूर्ण इंटरनेट डेटा के विशाल, अनियमित स्रोतों पर प्रशिक्षित स्व-शिक्षण तंत्रिका नेटवर्क के उपयोग को कम किया जाना चाहिए। [165]

पारदर्शिता की कमी

अधिकांश आधुनिक एआई एप्लिकेशन यह नहीं बता सकते कि वे किसी निर्णय पर कैसे पहुंचे। [166] गहरे तंत्रिका नेटवर्क में इनपुट और आउटपुट के बीच संबंधों की बड़ी मात्रा और परिणामी जटिलता एक विशेषज्ञ के लिए भी यह समझाना मुश्किल बना देती है कि उन्होंने अपने आउटपुट कैसे तैयार किए, जिससे वे एक ब्लैक बॉक्स बन गए । [167]

ऐसे कई मामले सामने आए हैं जहां एक मशीन लर्निंग प्रोग्राम ने कठोर परीक्षण पास किए, लेकिन फिर भी प्रोग्रामर ने जो इरादा किया था, उससे कुछ अलग सीखा। उदाहरण के लिए, जस्टिन को और रॉबर्टो नोवोआ ने एक ऐसी प्रणाली विकसित की जो चिकित्सा पेशेवरों की तुलना में त्वचा रोगों की बेहतर पहचान कर सकती है, हालांकि इसने रूलर वाली किसी भी छवि को "कैंसरयुक्त" के रूप में वर्गीकृत किया , क्योंकि घातक बीमारियों की तस्वीरों में आमतौर पर पैमाना दिखाने के लिए रूलर शामिल होता है। [168]2015 में रिच कारुआना द्वारा एक अधिक खतरनाक उदाहरण खोजा गया था: एक मशीन लर्निंग सिस्टम जिसने मृत्यु के जोखिम की सटीक भविष्यवाणी की थी, 65 वर्ष से अधिक उम्र के अस्थमा और सांस लेने में कठिनाई वाले रोगी को "कम जोखिम" के रूप में वर्गीकृत किया था। आगे के शोध से पता चला कि इस तरह के उच्च जोखिम वाले मामलों में, अस्पताल अधिक संसाधन आवंटित करेगा और रोगी के जीवन को बचाएगा, जिससे कार्यक्रम द्वारा मापा गया जोखिम कम हो जाएगा। [169] इस तरह की गलतियाँ तब स्पष्ट हो जाती हैं जब हमें पता चलता है कि कार्यक्रम किसी निर्णय पर कैसे पहुंचा है। स्पष्टीकरण के बिना, इन समस्याओं का पता तब तक नहीं लगाया जा सकता जब तक कि उन्होंने नुकसान न पहुँचाया हो।

दूसरा मुद्दा यह है कि जिन लोगों को एल्गोरिदम के निर्णय से नुकसान हुआ है, उन्हें स्पष्टीकरण का अधिकार है । उदाहरण के लिए, डॉक्टरों को अपने किसी भी निर्णय के पीछे के तर्क को स्पष्ट रूप से और पूरी तरह से समझाने की आवश्यकता होती है। [170] 2016 में यूरोपीय संघ के सामान्य डेटा संरक्षण विनियमन के शुरुआती मसौदे में एक स्पष्ट बयान शामिल था कि यह अधिकार मौजूद है। [ओ] उद्योग विशेषज्ञों ने कहा कि यह एक अनसुलझी समस्या है जिसका कोई समाधान नहीं दिख रहा है। नियामकों ने तर्क दिया कि फिर भी नुकसान वास्तविक है: यदि समस्या का कोई समाधान नहीं है, तो उपकरणों का उपयोग नहीं किया जाना चाहिए। [171]

DARPA ने इन समस्याओं को सुलझाने और हल करने के लिए 2014 में XAI ("एक्सप्लेनेबल आर्टिफिशियल इंटेलिजेंस") कार्यक्रम की स्थापना की । [172]

पारदर्शिता समस्या के कई संभावित समाधान हैं। मल्टीटास्क लर्निंग लक्ष्य वर्गीकरण के अलावा बड़ी संख्या में आउटपुट प्रदान करता है। ये अन्य आउटपुट डेवलपर्स को यह निष्कर्ष निकालने में मदद कर सकते हैं कि नेटवर्क ने क्या सीखा है। [173] डीकोनवोल्यूशन , डीपड्रीम और अन्य जेनरेटिव तरीके डेवलपर्स को यह देखने की अनुमति दे सकते हैं कि गहरे नेटवर्क की विभिन्न परतों ने क्या सीखा है और आउटपुट उत्पन्न करते हैं जो सुझाव दे सकते हैं कि नेटवर्क क्या सीख रहा है। [174] सुपरस्पार्स रैखिक पूर्णांक मॉडल वर्गीकरण के बजाय सबसे महत्वपूर्ण विशेषताओं की पहचान करने के लिए सीखने का उपयोग करते हैं। इन विशेषताओं को सरल रूप से जोड़ने से वर्गीकरण किया जा सकता है (अर्थात सीखने का उपयोग एक बनाने के लिए किया जाता है)।स्कोरिंग सिस्टम क्लासिफायर , जो पारदर्शी है)। [175]

बुरे अभिनेता और हथियारबंद एआई

घातक स्वायत्त हथियार एक ऐसी मशीन है जो मानव पर्यवेक्षण के बिना मानव लक्ष्यों का पता लगाती है, चुनती है और उन पर हमला करती है। [पी] 2015 तक, पचास से अधिक देशों में युद्धक्षेत्र रोबोटों पर शोध किए जाने की सूचना मिली थी। [177] इन हथियारों को कई कारणों से विशेष रूप से खतरनाक माना जाता है: यदि वे किसी निर्दोष व्यक्ति को मारते हैं तो यह स्पष्ट नहीं

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

है कि किसे जवाबदेह ठहराया जाना चाहिए , यह संभावना नहीं है कि वे विश्वसनीय रूप से लक्ष्य चुनेंगे, और, यदि बड़े पैमाने पर उत्पादित किया जाता है, तो वे संभावित रूप से हथियार हैं सामूहिक विनाश . [178] 2014 में, 30 देशों (चीन सहित) ने संयुक्त राष्ट्र के तहत स्वायत्त हथियारों पर प्रतिबंध का समर्थन किया ।कुछ पारंपरिक हथियारों पर कन्वेंशन , हालांकि संयुक्त राज्य अमेरिका और अन्य इससे सहमत नहीं थे। [179]

एआई कई उपकरण प्रदान करता है जो विशेष रूप से सत्तावादी सरकारों के लिए उपयोगी हैं: स्मार्ट स्पाइवेयर , चेहरे की पहचान और आवाज की पहचान व्यापक निगरानी की अनुमति देती है ; ऐसी निगरानी मशीन लर्निंग को राज्य के संभावित दुश्मनों को वर्गीकृत करने की अनुमति देती है और उन्हें छिपने से रोक सकती है; अनुशंसा प्रणालियाँ अधिकतम प्रभाव के लिए प्रचार और गलत सूचना को सटीक रूप से लक्षित कर सकती हैं ; डीपफेक गलत सूचना उत्पन्न करने में सहायता करता है; उन्नत AI सत्तावादी केंद्रीकृत निर्णय लेने में सक्षम हो सकता हैबाज़ार जैसी उदार और विकेन्द्रीकृत प्रणालियों के साथ अधिक प्रतिस्पर्धी । [180]

आतंकवादी, अपराधी और दुष्ट राज्य उन्नत डिजिटल युद्ध और घातक स्वायत्त हथियारों जैसे हथियारयुक्त एआई का उपयोग कर सकते हैं ।

मशीन-लर्निंग एआई कुछ ही घंटों में हजारों जहरीले अणुओं को डिजाइन करने में भी सक्षम है। [181]

तकनीकी बेरोजगारी

कृत्रिम बुद्धिमत्ता के विकास के शुरुआती दिनों से ही तर्क-वितर्क होते रहे हैं, उदाहरण के लिए वेइज़ेनबाम द्वारा सामने रखे गए तर्क , कि क्या कंप्यूटर द्वारा किए जा सकने वाले कार्य वास्तव में उनके द्वारा किए जाने चाहिए, कंप्यूटर और मनुष्यों के बीच और मात्रात्मक के बीच अंतर को देखते हुए गणना और गुणात्मक, मूल्य-आधारित निर्णय। [182]

अर्थशास्त्रियों ने अक्सर एआई से अतिरेक के जोखिमों पर प्रकाश डाला है, और पूर्ण रोजगार के लिए पर्याप्त सामाजिक नीति नहीं होने पर बेरोजगारी के बारे में अनुमान लगाया है। [183]

अतीत में, प्रौद्योगिकी ने कुल रोजगार को कम करने के बजाय बढ़ाने की प्रवृत्ति दिखाई है, लेकिन अर्थशास्त्री स्वीकार करते हैं कि एआई के साथ "हम अज्ञात क्षेत्र में हैं"। [184] अर्थशास्त्रियों के एक सर्वेक्षण में इस बात पर असहमति दिखाई गई कि क्या रोबोट और एआई के बढ़ते उपयोग से दीर्घकालिक बेरोजगारी में पर्याप्त वृद्धि होगी, लेकिन वे आम तौर पर सहमत हैं कि यदि उत्पादकता लाभ को पुनर्वितरित किया जाता है तो यह शुद्ध लाभ हो सकता है । [185] जोखिम का अनुमान अलग-अलग होता है; उदाहरण के लिए, 2010 के दशक में माइकल ओसबोर्न और कार्ल बेनेडिक्ट फ्रे ने अनुमान लगाया था कि 47% अमेरिकी नौकरियां संभावित स्वचालन के "उच्च जोखिम" में हैं, जबकि ओईसीडी की रिपोर्ट में केवल 9% अमेरिकी नौकरियों को "उच्च जोखिम" के रूप में वर्गीकृत किया गया है। [क्यू] [187]भविष्य में रोजगार के स्तर के बारे में अनुमान लगाने की पद्धति की साक्ष्य संबंधी आधार की कमी के रूप में आलोचना की गई है, और यह मानने के लिए कि प्रौद्योगिकी (सामाजिक नीति के बजाय) बेरोजगारी पैदा करती है (अतिरेक के विपरीत)। [183]

स्वचालन की पिछली लहरों के विपरीत, कृत्रिम बुद्धिमत्ता द्वारा कई मध्यम वर्ग की नौकरियाँ समाप्त हो सकती हैं; द इकोनॉमिस्ट ने 2015 में कहा था कि "यह चिंता कि एआई सफेदपोश नौकरियों के लिए वही कर सकता है जो औद्योगिक क्रांति के दौरान भाप की शक्ति ने नीलेपोशों के लिए किया था" "गंभीरता से लेने लायक है"। [188] अत्यधिक जोखिम वाली नौकरियों में पैरालीगल से लेकर फास्ट फूड पकाने वाले तक शामिल हैं, जबकि व्यक्तिगत स्वास्थ्य सेवा से लेकर पादरी तक देखभाल से संबंधित व्यवसायों के लिए नौकरी की मांग बढ़ने की संभावना है। [189]

अप्रैल 2023 में, यह बताया गया कि चीनी वीडियो गेम चित्रकारों की 70% नौकरियाँ जनरेटिव आर्टिफिशियल इंटेलिजेंस द्वारा समाप्त कर दी गई थीं। [190] [191]

कॉपीराइट

यथासंभव बड़े डेटासेट का लाभ उठाने के लिए, जेनरेटिव एआई को अक्सर बिना लाइसेंस वाले कॉपीराइट कार्यों पर प्रशिक्षित किया जाता है, जिसमें छवियों या कंप्यूटर कोड जैसे डोमेन शामिल हैं; फिर आउटपुट का उपयोग " उचित उपयोग " के तर्क के तहत किया जाता है। विशेषज्ञ इस बात पर असहमत हैं कि यह तर्क कितनी अच्छी तरह और किन परिस्थितियों में

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

अदालतों में टिकेगा; प्रासंगिक कारकों में "कॉपीराइट कार्य के उपयोग का उद्देश्य और चरित्र" और "कॉपीराइट कार्य के लिए संभावित बाजार पर प्रभाव" शामिल हो सकते हैं। [192]

नैतिक मशीनें और संरेखण

फ्रेंडली एआई ऐसी मशीनें हैं जिन्हें शुरू से ही जोखिमों को कम करने और मनुष्यों को लाभ पहुंचाने वाले विकल्प चुनने के लिए डिज़ाइन किया गया है। एलीएज़र युडकोव्स्की , जिन्होंने यह शब्द गढ़ा था, का तर्क है कि अनुकूल एआई विकसित करना एक उच्च अनुसंधान प्राथमिकता होनी चाहिए: इसके लिए बड़े निवेश की आवश्यकता हो सकती है और एआई के अस्तित्वगत जोखिम बनने से पहले इसे पूरा किया जाना चाहिए। [193]

बुद्धिमान मशीनों में नैतिक निर्णय लेने के लिए अपनी बुद्धि का उपयोग करने की क्षमता होती है। मशीन नैतिकता का क्षेत्र नैतिक दुविधाओं को हल करने के लिए मशीनों को नैतिक सिद्धांत और प्रक्रियाएं प्रदान करता है। [194] मशीन नैतिकता के क्षेत्र को कम्प्यूटेशनल नैतिकता भी कहा जाता है, [194] और इसकी स्थापना 2005 में एएएआई संगोष्ठी में की गई थी । [195]

अन्य दृष्टिकोणों में वेंडेल वैलाच के "कृत्रिम नैतिक एजेंट" [196] और स्टुअर्ट जे. रसेल के सिद्ध रूप से लाभकारी मशीनें विकसित करने के तीन सिद्धांत शामिल हैं। [197]

विनियमन

ओपनएआई के सीईओ सैम ऑल्टमैन ने अमेरिकी सीनेट उपसमिति, 2023 के समक्ष एआई विनियमन के बारे में गवाही दी

कृत्रिम बुद्धिमत्ता का विनियमन कृत्रिम बुद्धिमत्ता (एआई) को बढ़ावा देने और विनियमित करने के लिए सार्वजनिक क्षेत्र की नीतियों और कानूनों का विकास है; इसलिए यह एल्गोरिदम के व्यापक विनियमन से संबंधित है। [198] एआई के लिए विनियामक और नीतिगत परिदृश्य विश्व स्तर पर न्यायक्षेत्रों में एक उभरता हुआ मुद्दा है। [199] स्टैनफोर्ड में एआई इंडेक्स के अनुसार , 127 सर्वेक्षण देशों में पारित एआई-संबंधित कानूनों की वार्षिक संख्या 2016 में पारित एक से बढ़कर अकेले 2022 में 37 हो गई है। [200] [201] 2016 और 2020 के बीच, 30 से अधिक देशों ने एआई के लिए समर्पित रणनीतियाँ अपनाईं। [202] अधिकांश यूरोपीय संघ के सदस्य देशों ने राष्ट्रीय एआई रणनीतियाँ जारी की थीं, जैसे कि कनाडा, चीन, भारत, जापान, मॉरीशस, रूसी संघ, सऊदी अरब, संयुक्त अरब अमीरात, अमेरिका और वियतनाम। बांग्लादेश, मलेशिया और ट्यूनीशिया सहित अन्य लोग अपनी स्वयं की एआई रणनीति को विस्तृत करने की प्रक्रिया में थे। [202] आर्टिफिशियल इंटेलिजेंस पर वैश्विक साझेदारी जून 2020 में शुरू की गई थी, जिसमें प्रौद्योगिकी में जनता के विश्वास और विश्वास को सुनिश्चित करने के लिए मानव अधिकारों और लोकतांत्रिक मूल्यों के अनुसार एआई को विकसित करने की आवश्यकता बताई गई थी। [202] हेनरी किसिंजर , एरिक श्मिट और डैनियल हट्टेनलोचर ने नवंबर 2021 में एक संयुक्त बयान प्रकाशित किया जिसमें एआई को विनियमित करने के लिए एक सरकारी आयोग की मांग की गई। [203]2023 में, ओपनएआई नेताओं ने सुपरइंटेलिजेंस के प्रशासन के लिए सिफारिशें प्रकाशित कीं, जिनके बारे में उनका मानना है कि 10 साल से भी कम समय में ऐसा हो सकता है। [204]

2022 के इप्सोस सर्वेक्षण में , एआई के प्रति दृष्टिकोण देश के अनुसार काफी भिन्न था; 78% चीनी नागरिक, लेकिन केवल 35% अमेरिकी, इस बात से सहमत थे कि "एआई का उपयोग करने वाले उत्पादों और सेवाओं में कमियों की तुलना में अधिक लाभ हैं"। [200] 2023 रॉयटर्स /इप्सोस सर्वेक्षण में पाया गया कि 61% अमेरिकी सहमत हैं, और 22% असहमत हैं, कि एआई मानवता के लिए जोखिम पैदा करता है। [205] 2023 फॉक्स न्यूज पोल में, 35% अमेरिकियों ने इसे "बहुत महत्वपूर्ण" माना, और

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

अतिरिक्त 41% ने इसे संघीय सरकार के लिए एआई को विनियमित करने के लिए "कुछ हद तक महत्वपूर्ण" माना, जबकि 13% ने जवाब दिया "बहुत महत्वपूर्ण नहीं" और 8% ने जवाब दिया "बिल्कुल महत्वपूर्ण नहीं"। [206] [207]

इतिहास

यांत्रिक या "औपचारिक" तर्क का अध्ययन प्राचीन काल में दार्शनिकों और गणितज्ञों के साथ शुरू हुआ। तर्क का अध्ययन सीधे एलन ट्यूरिंग के गणना सिद्धांत की ओर ले गया, जिसने सुझाव दिया कि एक मशीन, "0" और "1" जैसे सरल प्रतीकों को फेरबदल करके, गणितीय कटौती और औपचारिक तर्क दोनों का अनुकरण कर सकती है, जिसे चर्च के रूप में जाना जाता है । -ट्यूरिंग थीसिस . [208] इसने, साइबरनेटिक्स और सूचना सिद्धांत में समवर्ती खोजों के साथ , शोधकर्ताओं को "इलेक्ट्रॉनिक मस्तिष्क" के निर्माण की संभावना पर विचार करने के लिए प्रेरित किया।1943 में ट्यूरिंग-पूर्ण "कृत्रिम न्यूरॉन्स" के लिए डिज़ाइन। [211]

एआई अनुसंधान के क्षेत्र की स्थापना 1956 में डार्टमाउथ कॉलेज में एक कार्यशाला में की गई थी । [एस] [2] उपस्थित लोग 1960 के दशक में एआई अनुसंधान के नेता बन गए। [टी] उन्होंने और उनके छात्रों ने ऐसे कार्यक्रम तैयार किए जिन्हें प्रेस ने "आश्चर्यजनक" बताया: [यू] कंप्यूटर चेकर्स रणनीतियों को सीख रहे थे , बीजगणित में शब्द समस्याओं को हल कर रहे थे, तार्किक प्रमेयों को साबित कर रहे थे और अंग्रेजी बोल रहे थे। [v] [3]

1960 के दशक के मध्य तक, अमेरिका में अनुसंधान को रक्षा विभाग द्वारा भारी वित्त पोषित किया गया था [215] और दुनिया भर में प्रयोगशालाएँ स्थापित की गई थीं। [216] हर्बर्ट साइमन ने भविष्यवाणी की थी, "बीस वर्षों के भीतर मशीनें कोई भी काम करने में सक्षम होंगी जो एक आदमी कर सकता है"। [217] मार्विन मिंस्की ने सहमति व्यक्त करते हुए लिखा, "एक पीढ़ी के भीतर... 'कृत्रिम बुद्धिमत्ता' बनाने की समस्या काफी हद तक हल हो जाएगी"। [218]

हालाँकि, उन्होंने समस्या की कठिनाई को कम करके आंका था। [w] सर जेम्स लाइटहिल की आलोचना [220] और अधिक उत्पादक परियोजनाओं को वित्त पोषित करने के लिए अमेरिकी कांग्रेस के चल रहे दबाव के जवाब में अमेरिकी और ब्रिटिश दोनों सरकारों ने खोजपूर्ण अनुसंधान बंद कर दिया । मिन्स्की और पैपर्ट की पुस्तक पर्सेप्ट्रॉन को यह साबित करने के रूप में समझा गया कि कृत्रिम तंत्रिका नेटवर्क दृष्टिकोण वास्तविक दुनिया के कार्यों को हल करने के लिए कभी भी उपयोगी नहीं होगा, इस प्रकार दृष्टिकोण पूरी तरह से बदनाम हो गया। [221] इसके बाद " एआई विंटर " आया, एक ऐसा दौर जब एआई परियोजनाओं के लिए धन प्राप्त करना कठिन था।[5]

1980 के दशक की शुरुआत में, विशेषज्ञ प्रणालियों की व्यावसायिक सफलता से एआई अनुसंधान को पुनर्जीवित किया गया , [222] एआई कार्यक्रम का एक रूप जिसने मानव विशेषज्ञों के ज्ञान और विश्लेषणात्मक कौशल का अनुकरण किया। 1985 तक, AI का बाज़ार एक अरब डॉलर से अधिक तक पहुँच गया था। उसी समय, जापान की पांचवीं पीढ़ी के कंप्यूटर प्रोजेक्ट ने अमेरिकी और ब्रिटिश सरकारों को अकादमिक अनुसंधान के लिए धन बहाल करने के लिए प्रेरित किया । [4] हालाँकि, 1987 में लिस्प मशीन बाजार के पतन के साथ शुरुआत करते हुए , एआई एक बार फिर बदनाम हो गया और दूसरी, लंबे समय तक चलने वाली सर्दी शुरू हो गई। [6]

कई शोधकर्ताओं को संदेह होने लगा कि वर्तमान प्रथाएं मानव अनुभूति की सभी प्रक्रियाओं, विशेष रूप से धारणा , रोबोटिक्स, सीखने और पैटर्न पहचान की नकल करने में सक्षम होंगी । [223] कई शोधकर्ताओं ने "उप-प्रतीकात्मक" दृष्टिकोणों पर गौर करना शुरू किया। [224] रॉडनी ब्रूक्स जैसे रोबोटिक्स शोधकर्ताओं ने सामान्य रूप से "प्रतिनिधित्व" को खारिज कर दिया और सीधे इंजीनियरिंग मशीनों पर ध्यान केंद्रित किया जो चलती हैं और जीवित रहती हैं। [x] . जुडिया पर्ल , लोफ़्टी ज़ादेह और अन्य ने ऐसे तरीके विकसित किए जो सटीक तर्क के बजाय उचित अनुमान लगाकर अधूरी और अनिश्चित जानकारी को संभालते थे। [83][229] लेकिन सबसे महत्वपूर्ण विकास जेफ्री हिंटन और अन्य द्वारा तंत्रिका नेटवर्क अनुसंधानसहित " कनेक्शनवाद " का पुनरुद्धार था। [230] 1990 में, यान लेकुन ने सफलतापूर्वक दिखाया कि कन्वेन्शनल न्यूरल नेटवर्क हस्तलिखित अंकों को पहचान सकते हैं, जो न्यूरल नेटवर्क के कई सफल अनुप्रयोगों में से पहला था। [231]

एआई ने औपचारिक गणितीय तरीकों का फायदा उठाकर और विशिष्ट समस्याओं के विशिष्ट समाधान ढूंढकर 1990 के दशक के अंत और 21वीं सदी की शुरुआत में धीरे-धीरे अपनी प्रतिष्ठा बहाल की। इस " संकीर्ण " और "औपचारिक" फोकस ने

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

शोधकर्ताओं को सत्यापन योग्य परिणाम उत्पन्न करने और अन्य क्षेत्रों (जैसे सांख्यिकी , अर्थशास्त्र और गणित) के साथ सहयोग करने की अनुमति दी। [232] 2000 तक, एआई शोधकर्ताओं द्वारा विकसित समाधानों का व्यापक रूप से उपयोग किया जा रहा था, हालांकि 1990 के दशक में उन्हें शायद ही कभी "कृत्रिम बुद्धिमत्ता" के रूप में वर्णित किया गया था। [233]

कई अकादमिक शोधकर्ता चिंतित हो गए कि एआई अब बहुमुखी, पूरी तरह से बुद्धिमान मशीनें बनाने के मूल लक्ष्य का पीछा नहीं कर रहा है। 2002 के आसपास शुरुआत करते हुए, उन्होंने कृत्रिम सामान्य बुद्धिमत्ता (या "एजीआई") के उपक्षेत्र की स्थापना की , जिसमें 2010 तक कई अच्छी तरह से वित्त पोषित संस्थान थे। [8]

2012 में गहन शिक्षा उद्योग के बेंचमार्क पर हावी होने लगी और इसे पूरे क्षेत्र में अपनाया गया। [7] कई विशिष्ट कार्यों के लिए, अन्य तरीकों को छोड़ दिया गया। [y] डीप लर्निंग की सफलता हार्डवेयर सुधार (तेज कंप्यूटर , [235] ग्राफिक्स प्रोसेसिंग यूनिट , क्लाउड कंप्यूटिंग [236]) और बड़ी मात्रा में डेटा तक पहुंच [237] (क्यूरेटेड डेटासेट सहित, [236] जैसे कि इमेजनेट) दोनों पर आधारित थी।).

डीप लर्निंग की सफलता से एआई में रुचि और फंडिंग में भारी वृद्धि हुई। [z] वर्ष 2015-2019 में मशीन लर्निंग अनुसंधान की मात्रा (कुल प्रकाशनों द्वारा मापी गई) 50% की वृद्धि हुई, [202] और डब्ल्यूआईपीओ ने बताया कि एआई पेटेंट आवेदनों की संख्या के मामले में सबसे उभरती हुई तकनीक थी और दी गई थी। पेटेंट [238] 'एआई इम्पैक्ट्स' के अनुसार, अकेले अमेरिका में 2022 के आसपास "एआई" में सालाना लगभग 50 बिलियन डॉलर का निवेश किया गया था और लगभग 20% नए अमेरिकी कंप्यूटर विज्ञान पीएचडी स्नातकों ने "एआई" में विशेषज्ञता हासिल की है; [239] 2022 में लगभग 800,000 "एआई"-संबंधित अमेरिकी नौकरियों के अवसर मौजूद थे। [240]

2016 में, मशीन लर्निंग सम्मेलनों में निष्पक्षता और प्रौद्योगिकी के दुरुपयोग के मुद्दों को केंद्र स्तर पर ले जाया गया, प्रकाशनों में भारी वृद्धि हुई, धन उपलब्ध हो गया, और कई शोधकर्ताओं ने इन मुद्दों पर अपने करियर को फिर से केंद्रित किया। संरेखण समस्या अकादमिक अध्ययन का एक गंभीर क्षेत्र बन गई। [241]

दर्शन

कृत्रिम बुद्धिमत्ता को परिभाषित करना

एलन ट्यूरिंग ने 1950 में लिखा था "मैं इस प्रश्न पर विचार करने का प्रस्ताव करता हूं कि 'क्या मशीनें सोच सकती हैं'?" [242] उन्होंने इस सवाल को बदलने की सलाह दी कि क्या कोई मशीन "सोचती है" से "मशीनरी के लिए बुद्धिमान व्यवहार दिखाना संभव है या नहीं"। [242] उन्होंने ट्यूरिंग परीक्षण तैयार किया, जो मानव वार्तालाप का अनुकरण करने के लिए एक मशीन की क्षमता को मापता है। [243] चूंकि हम केवल मशीन के व्यवहार का निरीक्षण कर सकते हैं, इससे कोई फर्क नहीं पड़ता कि वह "वास्तव में" सोच रही है या सचमुच उसके पास "दिमाग" है। ट्यूरिंग नोट करते हैं कि हम अन्य लोगों के बारे में ये बातें निर्धारित नहीं कर सकते हैं [एए] लेकिन "एक विनम्र सम्मेलन होना सामान्य बात है जिसके बारे में हर कोई सोचता है" [244]

रसेल और नॉरविग ट्यूरिंग से सहमत हैं कि एआई को "अभिनय" के संदर्भ में परिभाषित किया जाना चाहिए न कि "सोचने" के संदर्भ में। [245] हालांकि, वे महत्वपूर्ण हैं कि परीक्षण मशीनों की तुलना लोगों से करता है । " वैमानिकी इंजीनियरिंग पाठ," उन्होंने लिखा, "अपने क्षेत्र के लक्ष्य को 'ऐसी मशीनें बनाने के रूप में परिभाषित नहीं करते हैं जो बिल्कुल कबूतरों की तरह उड़ती हैं ताकि वे अन्य कबूतरों को बेवकूफ बना सकें। ' " [246] एआई के संस्थापक जॉन मैक्कार्थी सहमत हुए, उन्होंने लिखा कि "कृत्रिम बुद्धि, परिभाषा के अनुसार, मानव बुद्धि का अनुकरण नहीं है"। [247]

मैक्कार्थी ने बुद्धिमत्ता को "दुनिया में लक्ष्य प्राप्त करने की क्षमता का कम्प्यूटेशनल हिस्सा" के रूप में परिभाषित किया है। [248] एआई के एक अन्य संस्थापक, मार्विन मिंस्की ने इसे इसी तरह "कठिन समस्याओं को हल करने की क्षमता" के रूप में परिभाषित किया है। [249] ये परिभाषाएँ बुद्धिमत्ता को अच्छी तरह से परिभाषित समस्याओं के साथ अच्छी तरह से परिभाषित समाधानों के संदर्भ में देखती हैं, जहाँ समस्या की कठिनाई और कार्यक्रम का प्रदर्शन दोनों मशीन की "बुद्धिमत्ता" के प्रत्यक्ष उपाय हैं - और कोई अन्य दार्शनिक चर्चा नहीं है आवश्यक है, या संभव भी नहीं हो सकता है।

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

एक अन्य परिभाषा Google [250] [बेहतर स्रोत की आवश्यकता] द्वारा अपनाई गई है , जो AI के क्षेत्र में एक प्रमुख व्यवसायी है। यह परिभाषा बुद्धिमत्ता की अभिव्यक्ति के रूप में सूचना को संश्लेषित करने की सिस्टम की क्षमता को निर्धारित करती है, ठीक उसी तरह जैसे इसे जैविक बुद्धिमत्ता में परिभाषित किया गया है।

एआई के दृष्टिकोण का मूल्यांकन

किसी स्थापित एकीकृत सिद्धांत या प्रतिमान ने इसके अधिकांश इतिहास में एआई अनुसंधान का मार्गदर्शन नहीं किया है। [एबी] 2010 के दशक में सांख्यिकीय मशीन लर्निंग की अभूतपूर्व सफलता ने अन्य सभी दृष्टिकोणों को पीछे छोड़ दिया (इतना कि कुछ स्रोत, विशेष रूप से व्यापार जगत में, "कृत्रिम बुद्धिमत्ता" शब्द का उपयोग "तंत्रिका नेटवर्क के साथ मशीन लर्निंग" के लिए करते हैं)। यह दृष्टिकोण अधिकतर उप-प्रतीकात्मक , नरम और संकीर्ण है (नीचे देखें)। आलोचकों का तर्क है कि एआई शोधकर्ताओं की भावी पीढ़ियों को इन सवालों पर दोबारा विचार करना पड़ सकता है।

प्रतीकात्मक एआई और इसकी सीमाएँ

प्रतीकात्मक AI (या " GOFAI ") [252] ने उच्च-स्तरीय सचेत तर्क का अनुकरण किया है जिसका उपयोग लोग पहेलियाँ हल करते समय, कानूनी तर्क व्यक्त करते समय और गणित करते समय करते हैं। वे बीजगणित या बुद्धि परीक्षण जैसे "बुद्धिमान" कार्यों में अत्यधिक सफल रहे। 1960 के दशक में, नेवेल और साइमन ने भौतिक प्रतीक प्रणाली परिकल्पना का प्रस्ताव रखा : "एक भौतिक प्रतीक प्रणाली में सामान्य बुद्धिमान कार्रवाई के आवश्यक और पर्याप्त साधन होते हैं।" [253]

हालाँकि, प्रतीकात्मक दृष्टिकोण कई कार्यों में विफल रहा, जिन्हें मनुष्य आसानी से हल कर लेते हैं, जैसे सीखना, किसी वस्तु को पहचानना या सामान्य ज्ञान तर्क। मोरावेक का विरोधाभास यह खोज है कि एआई के लिए उच्च-स्तरीय "बुद्धिमान" कार्य आसान थे, लेकिन निम्न-स्तरीय "सहज" कार्य बेहद कठिन थे। [254] दार्शनिक ह्यूबर्ट ड्रेफस ने 1960 के दशक से तर्क दिया था कि मानव विशेषज्ञता सचेत प्रतीक हेरफेर के बजाय अचेतन वृत्ति पर निर्भर करती है, और स्पष्ट प्रतीकात्मक ज्ञान के बजाय स्थिति के लिए "महसूस" करने पर निर्भर करती है। [255] हालाँकि जब उनके तर्कों को पहली बार प्रस्तुत किया गया था तो उनका उपहास किया गया था और उन्हें नजरअंदाज कर दिया गया था, अंततः, एआई अनुसंधान सहमत हुआ। [एसी] [13]

समस्या हल नहीं हुई है: उप-प्रतीकात्मक तर्क कई ऐसी ही गूढ़ गलतियाँ कर सकता है जो मानव अंतर्ज्ञान करता है, जैसे कि एल्गोरिथम पूर्वाग्रह । नोम चॉम्स्की जैसे आलोचकों का तर्क है कि सामान्य बुद्धिमत्ता प्राप्त करने के लिए प्रतीकात्मक एआई में अनुसंधान जारी रखना अभी भी आवश्यक होगा, [257] [258] आंशिक रूप से क्योंकि उप-प्रतीकात्मक एआई समझाने योग्य एआई से एक कदम दूर है : यह समझना मुश्किल या असंभव हो सकता है कि ऐसा क्यों है एक आधुनिक सांख्यिकीय एआई कार्यक्रम ने एक विशेष निर्णय लिया। न्यूरो-प्रतीकात्मक कृत्रिम बुद्धिमत्ता का उभरता हुआ क्षेत्र दो दृष्टिकोणों को पाटने का प्रयास करता है।

साफ़-सुथरा बनाम मैला-कुचैला

"नीट्स" को उम्मीद है कि बुद्धिमान व्यवहार को सरल, सुरुचिपूर्ण सिद्धांतों (जैसे तर्क , अनुकूलन , या तंत्रिका नेटवर्क) का उपयोग करके वर्णित किया गया है। "स्क्रफ़ीज़" उम्मीद करते हैं कि इसके लिए आवश्यक रूप से बड़ी संख्या में असंबंधित समस्याओं को हल करने की आवश्यकता है। नीट सैद्धांतिक कठोरता के साथ अपने कार्यक्रमों का बचाव करते हैं, स्क्रूफी मुख्य रूप से वृद्धिशील परीक्षण पर भरोसा करते हैं यह देखने के लिए कि क्या वे काम करते हैं। इस मुद्दे पर 70 और 80 के दशक में सक्रिय रूप से चर्चा की गई थी, [259] लेकिन अंततः इसे अप्रासंगिक के रूप में देखा गया। आधुनिक AI में दोनों के तत्व हैं।

सॉफ्ट बनाम हार्ड कंप्यूटिंग

कई महत्वपूर्ण समस्याओं के लिए एक सिद्ध रूप से सही या इष्टतम समाधान ढूंढना कठिन है। [12] सॉफ्ट कंप्यूटिंग तकनीकों का एक सेट है, जिसमें जेनेटिक एल्गोरिदम , फ़ज़ी लॉजिक और न्यूरल नेटवर्क शामिल हैं, जो अशुद्धि, अनिश्चितता, आंशिक सत्य और सन्निकटन के प्रति सहनशील हैं। सॉफ्ट कंप्यूटिंग की शुरुआत 80 के दशक के अंत में हुई थी और 21वीं सदी में सबसे सफल एआई प्रोग्राम न्यूरल नेटवर्क के साथ सॉफ्ट कंप्यूटिंग के उदाहरण हैं।

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

संकीर्ण बनाम सामान्य एआई

एआई शोधकर्ता इस बात पर विभाजित हैं कि क्या कृत्रिम सामान्य बुद्धिमत्ता और सुपरइंटेलिजेंस के लक्ष्यों को सीधे आगे बढ़ाया जाए या यथासंभव अधिक से अधिक विशिष्ट समस्याओं (संकीर्ण एआई) को हल किया जाए, इस उम्मीद में कि ये समाधान अप्रत्यक्ष रूप से क्षेत्र के दीर्घकालिक लक्ष्यों की ओर ले जाएंगे। [260] [261] सामान्य बुद्धिमत्ता को परिभाषित करना और मापना कठिन है, और आधुनिक एआई को विशिष्ट समाधानों के साथ विशिष्ट समस्याओं पर ध्यान केंद्रित करके अधिक सत्यापन योग्य सफलता मिली है। कृत्रिम सामान्य बुद्धि का प्रायोगिक उप-क्षेत्र विशेष रूप से इस क्षेत्र का अध्ययन करता है।

मशीन चेतना, भावना और मन

मन का दर्शन यह नहीं जानता कि क्या किसी मशीन में मन , चेतना और मानसिक अवस्थाएं उसी तरह हो सकती हैं, जैसे मनुष्य में होती हैं। यह मुद्दा मशीन के बाहरी व्यवहार के बजाय उसके आंतरिक अनुभवों पर विचार करता है। मुख्यधारा एआई अनुसंधान इस मुद्दे को अप्रासंगिक मानता है क्योंकि यह क्षेत्र के लक्ष्यों को प्रभावित नहीं करता है: ऐसी मशीनें बनाना जो बुद्धि का उपयोग करके समस्याओं को हल कर सकें। रसेल और नॉरविग कहते हैं कि "[टी] वह एक मशीन को इंसानों की तरह जागरूक बनाने की अतिरिक्त परियोजना है जिसे लेने के लिए हम सुसज्जित नहीं हैं।" [262]हालाँकि, यह प्रश्न मन के दर्शन का केंद्र बन गया है। यह आम तौर पर कथा साहित्य में कृत्रिम बुद्धिमत्ता के मुद्दे पर केंद्रीय प्रश्न भी है ।

चेतना

डेविड चाल्मर्स ने मन को समझने में दो समस्याओं की पहचान की, जिन्हें उन्होंने चेतना की "कठिन" और "आसान" समस्याओं का नाम दिया। [263] आसान समस्या यह समझना है कि मस्तिष्क कैसे संकेतों को संसाधित करता है, योजना बनाता है और व्यवहार को नियंत्रित करता है। कठिन समस्या यह समझाना है कि यह कैसा महसूस होता है या इसे किसी भी चीज़ की तरह क्यों महसूस होना चाहिए, यह मानते हुए कि हम यह सोचने में सही हैं कि यह वास्तव में कुछ ऐसा महसूस करता है (डेनेट की चेतना भ्रमवाद कहता है कि यह एक भ्रम है)। हालाँकि, मानव सूचना प्रसंस्करण को समझाना आसान है, मानव व्यक्तिपरक अनुभवसमझाना कठिन है. उदाहरण के लिए, एक रंग-अंध व्यक्ति की कल्पना करना आसान है जिसने यह पहचानना सीख लिया है कि उसके दृश्य क्षेत्र में कौन सी वस्तुएं लाल हैं, लेकिन यह स्पष्ट नहीं है कि उस व्यक्ति को यह जानने के लिए क्या आवश्यक होगा कि लाल कैसा दिखता है । [264]

कम्प्यूटेशनलिज्म और कार्यात्मकता

कम्प्यूटेशनलिज्म मन के दर्शन में यह स्थिति है कि मानव मन एक सूचना प्रसंस्करण प्रणाली है और सोच कंप्यूटिंग का एक रूप है। कम्प्यूटेशनलवाद का तर्क है कि मन और शरीर के बीच का संबंध सॉफ्टवेयर और हार्डवेयर के बीच के संबंध के समान या समान है और इस प्रकार मन -शरीर की समस्या का समाधान हो सकता है । यह दार्शनिक स्थिति 1960 के दशक में एआई शोधकर्ताओं और संज्ञानात्मक वैज्ञानिकों के काम से प्रेरित थी और मूल रूप से दार्शनिक जेरी फोडर और हिलेरी पटनम द्वारा प्रस्तावित की गई थी । [265]

दार्शनिक जॉन सियरल ने इस स्थिति को " मजबूत एआई " के रूप में वर्णित किया: "सही इनपुट और आउटपुट के साथ उचित रूप से प्रोग्राम किए गए कंप्यूटर में ठीक उसी तरह एक दिमाग होगा जैसे इंसानों के पास होता है।" [विज्ञापन] सियरल ने अपने चीनी कमरे के तर्क के साथ इस दावे का खंडन किया है, जो यह दिखाने का प्रयास करता है कि, भले ही एक मशीन मानव व्यवहार का पूरी तरह से अनुकरण करती है, फिर भी यह मानने का कोई कारण नहीं है कि इसमें दिमाग भी है। [269]

रोबोट अधिकार

यदि किसी मशीन में दिमाग और व्यक्तिपरक अनुभव है, तो उसमें संवेदना (महसूस करने की क्षमता) भी हो सकती है, और यदि ऐसा है तो उसे नुकसान भी हो सकता है ; यह तर्क दिया गया है कि इससे उसे कुछ अधिकार प्राप्त हो सकते हैं। [270] कोई भी काल्पनिक रोबोट अधिकार पशु अधिकारों और मानवाधिकारों के स्पेक्ट्रम पर आधारित होगा । [271] इस मुद्दे पर सदियों से कल्पना में विचार किया गया है, [272] और अब, उदाहरण के लिए, कैलिफोर्निया के इंस्टीट्यूट फॉर द फ्यूचर द्वारा इस पर विचार किया जा रहा है ; हालाँकि, आलोचकों का तर्क है कि चर्चा समय से पहले हुई है। [273]

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

भविष्य

अधीक्षण और विलक्षणता

सुपरइंटेलिजेंस एक काल्पनिक एजेंट है जिसके पास सबसे प्रतिभाशाली और सबसे प्रतिभाशाली मानव दिमाग से कहीं अधिक बुद्धिमत्ता होगी। [261]

यदि कृत्रिम सामान्य बुद्धिमत्ता में अनुसंधान से पर्याप्त बुद्धिमान सॉफ़्टवेयर का उत्पादन होता है, तो यह स्वयं को पुन: प्रोग्राम करने और सुधारने में सक्षम हो सकता है। बेहतर सॉफ्टवेयर खुद को बेहतर बनाने में और भी बेहतर होगा, जिससे आईजे गुड ने " खुफिया विस्फोट " कहा और वर्नोर विंग ने " विलक्षणता " कहा । [274] हालांकि, अधिकांश प्रौद्योगिकियां अनिश्चित काल तक तेजी से सुधार नहीं करती हैं, बल्कि एस-वक्र का पालन करती हैं, जब वे प्रौद्योगिकी जो कर सकती हैं उसकी भौतिक सीमा तक पहुंचने पर धीमी हो जाती हैं। [275] उदाहरण के लिए, परिवहन पर विचार करें: 1830 से 1970 तक इसमें तेजी से सुधार हुआ, लेकिन भौतिक सीमा तक पहुंचने पर यह प्रवृत्ति अचानक बंद हो गई।

अस्तित्वगत जोखिम

यह तर्क दिया गया है कि एआई इतना शक्तिशाली हो जाएगा कि मानवता अपरिवर्तनीय रूप से इस पर नियंत्रण खो देगी। जैसा कि भौतिक विज्ञानी स्टीफ़न हॉकिंग कहते हैं, यह " मानव जाति के अंत का कारण बन सकता है "। [276] यह परिदृश्य विज्ञान कथाओं में आम रहा है, जब एक कंप्यूटर या रोबोट अचानक मानव जैसी "आत्म-जागरूकता" (या "संवेदना" या "चेतना") विकसित करता है और एक द्वेषपूर्ण चरित्र बन जाता है। [एई] ये विज्ञान-फाई परिदृश्य कई मायनों में भ्रामक हैं।

सबसे पहले, एआई को अस्तित्वगत जोखिम के लिए मानव जैसी "भावना" की आवश्यकता नहीं है। आधुनिक एआई कार्यक्रमों को विशिष्ट लक्ष्य दिए जाते हैं और उन्हें प्राप्त करने के लिए सीखने और बुद्धि का उपयोग किया जाता है। दार्शनिक निक बोस्ट्रोम ने तर्क दिया कि यदि कोई पर्याप्त रूप से शक्तिशाली एआई को लगभग कोई भी लक्ष्य देता है, तो वह इसे प्राप्त करने के लिए मानवता को नष्ट करने का विकल्प चुन सकता है (उन्होंने पेपरक्लिप फैक्ट्री प्रबंधक के उदाहरण का उपयोग किया)। [278] स्टुअर्ट रसेल घरेलू रोबोट का उदाहरण देते हैं जो अपने मालिक को मारने का तरीका खोजने की कोशिश करता है ताकि उसका प्लग बंद न हो, और तर्क देता है कि "यदि आप मर गए तो आप कॉफी नहीं ला सकते।" [279] मानवता के लिए सुरक्षित रहने के लिए, एक सुपरइंटेलिजेंस को वास्तव में संरेखित करना होगामानवता की नैतिकता और मूल्यों के साथ ताकि यह "मौलिक रूप से हमारे पक्ष में" हो। [280]

दूसरा, युवल नूह हरारी का तर्क है कि एआई को अस्तित्वगत जोखिम पैदा करने के लिए रोबोट बॉडी या भौतिक नियंत्रण की आवश्यकता नहीं है। सभ्यता के आवश्यक अंग भौतिक नहीं हैं। विचारधारा , कानून , सरकार , पैसा और अर्थव्यवस्था जैसी चीजें भाषा से बनती हैं ; वे अस्तित्व में हैं क्योंकि ऐसी कहानियाँ हैं जिन पर अरबों लोग विश्वास करते हैं। गलत सूचना की वर्तमान व्यापकता से पता चलता है कि एआई लोगों को किसी भी बात पर विश्वास करने के लिए भाषा का उपयोग कर सकता है, यहां तक कि विनाशकारी कार्रवाई करने के लिए भी। [281]

विशेषज्ञों और उद्योग के अंदरूनी सूत्रों के बीच राय मिश्रित है, जिसमें बड़े पैमाने पर अंश अंततः सुपरइंटेलिजेंट एआई के जोखिम से चिंतित और असंबद्ध दोनों हैं। [282] स्टीफन हॉकिंग , बिल गेट्स , एलोन मस्क जैसी हस्तियों ने एआई से अस्तित्व संबंधी खतरे के बारे में चिंता व्यक्त की है। [283] 2010 की शुरुआत में, विशेषज्ञों ने तर्क दिया कि भविष्य में अनुसंधान की आवश्यकता के लिए जोखिम बहुत दूर हैं, या मनुष्य एक सुपरइंटेलिजेंट मशीन के नजरिए से मूल्यवान होंगे। [284] हालांकि, 2016 के बाद, वर्तमान और भविष्य के जोखिमों और संभावित समाधानों का अध्ययन अनुसंधान का एक गंभीर क्षेत्र बन गया। [241] 2023 में, जेफ्री हिंटन सहित एआई अग्रणी, योशुआ बेंगियो , डेमिस हसाबिस और सैम ऑल्टमैन ने संयुक्त बयान जारी किया कि "महामारी और परमाणु युद्ध जैसे अन्य सामाजिक-स्तर के जोखिमों के साथ-साथ एआई से विलुप्त होने के जोखिम को कम करना एक वैश्विक प्राथमिकता होनी चाहिए"; यान लेकुन जैसे कुछ अन्य लोग इसे निराधार मानते हैं। [285]

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

ट्रांसह्युमेनिज़म

रोबोट डिजाइनर हंस मोरवेक , साइबरनेटिसिस्ट केविन वारविक और आविष्कारक रे कुर्ज़वील ने भविष्यवाणी की है कि भविष्य में मनुष्य और मशीनें साइबरबॉर्ग में विलीन हो जाएंगी जो कि दोनों की तुलना में अधिक सक्षम और शक्तिशाली हैं। यह विचार, जिसे ट्रांसह्यूमनिज़्म कहा जाता है, की जड़ें एल्डस हक्सले और रॉबर्ट एटिंगर में हैं । [286]

एडवर्ड फ्रेडकिन का तर्क है कि "कृत्रिम बुद्धिमत्ता विकास का अगला चरण है", यह विचार सबसे पहले 1863 में सैमुअल बटलर की " डार्विन अमंग द मशीन्स " द्वारा प्रस्तावित किया गया था, और जॉर्ज डायसन ने इसी नाम की अपनी पुस्तक में इसका विस्तार किया था। 1998. [287]

निष्कर्ष

"रोबोट" शब्द का प्रयोग कारेल कैपेक ने अपने 1921 के नाटक आरयूआर में किया था , जिसका शीर्षक "रोसुम्स यूनिवर्सल रोबोट्स" है।

विचार-सक्षम कृत्रिम प्राणी प्राचीन काल से कहानी कहने के उपकरणों के रूप में प्रकट हुए हैं, [288] और विज्ञान कथा में एक स्थायी विषय रहे हैं । [289]

इन कार्यों में एक आम बात मैरी शेली के फ्रेंकस्टीन के साथ शुरू हुई , जहां एक मानव रचना अपने मालिकों के लिए खतरा बन जाती है। इसमें आर्थर सी. क्लार्क और स्टेनली कुब्रिक की 2001: ए स्पेस ओडिसी (दोनों 1968), एचएएल 9000 के साथ, डिस्कवरी वन स्पेसशिप के प्रभारी जानलेवा कंप्यूटर , साथ ही द टर्मिनेटर (1984) और द मैट्रिक्स (1999) जैसे काम शामिल हैं।). इसके विपरीत, द डे द अर्थ स्टूड स्टिल (1951) के गोर्ट और एलियंस (1986) के बिशप जैसे दुर्लभ वफादार रोबोट लोकप्रिय संस्कृति में कम प्रमुख हैं। [290]

आइजैक असिमोव ने कई किताबों और कहानियों में रोबोटिक्स के तीन नियमों को पेश किया , विशेष रूप से इसी नाम के एक सुपर-इंटेलिजेंट कंप्यूटर के बारे में "मल्टीवैक" श्रृंखला। असिमोव के नियम अक्सर मशीनी नैतिकता की चर्चा के दौरान सामने आते हैं; [291] जबकि लगभग सभी कृत्रिम बुद्धिमत्ता शोधकर्ता लोकप्रिय संस्कृति के माध्यम से असिमोव के नियमों से परिचित हैं, वे आम तौर पर कई कारणों से कानूनों को बेकार मानते हैं, जिनमें से एक उनकी अस्पष्टता है। [292]

कई कार्य एआई का उपयोग करते हुए हमें उस मूलभूत प्रश्न का सामना करने के लिए मजबूर करते हैं जो हमें मानव बनाता है, हमें कृत्रिम प्राणी दिखाता है जिसमें महसूस करने और इस प्रकार पीड़ित होने की क्षमता होती है। यह कारेल कैपेक के आरयूआर , एआई आर्टिफिशियल इंटेलिजेंस और एक्स मशीना फिल्मों के साथ-साथ उपन्यास डू एंड्रॉइड्स ड्रीम ऑफ इलेक्ट्रिक शीप में भी दिखाई देता है। , फिलिप के. डिक द्वारा । डिक इस विचार पर विचार करते हैं कि मानव व्यक्तिपरकता की हमारी समझ कृत्रिम बुद्धिमत्ता से निर्मित प्रौद्योगिकी द्वारा बदल दी गई है। [293]

प्रतिक्रिया दें संदर्भ

1. गूगल (2016) ।
2. ^ऊपर जायें:ए बी डार्टमाउथ कार्यशाला:
   - रसेल और नॉरविग (2021 , पृष्ठ 18)

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

- मैककॉर्डक (2004 , पृ. 111-136)
- एनआरसी (1999 , पृ. 200-201)

प्रस्ताव:

- मैक्कार्थी एट अल. (1955)

3. ^ऊपर जायें:[ए बी] 60 के दशक के सफल कार्यक्रम:
   - मैककॉर्डक (2004 , पृ. 243-252)
   - क्रेवियर (1993 , पृ. 52-107)
   - मोरावेक (1988 , पृ. 9)
   - रसेल और नॉरविग (2021 , पीपी. 19-21)
4. ^ऊपर जायें:[ए बी] 80 के दशक की शुरुआत में फंडिंग पहल:फिफ्थ जेनरेशन प्रोजेक्ट(जापान),एल्वे(यूके),माइक्रोइलेक्ट्रॉनिक्स एंड कंप्यूटर टेक्नोलॉजी कॉरपोरेशन(यूएस),स्ट्रैटेजिक कंप्यूटिंग इनिशिएटिव(यूएस):
   - मैककॉर्डक (2004 , पृ. 426-441)
   - क्रेवियर (1993 , पृ. 161-162, 197-203, 211, 240)
   - रसेल और नॉरविग (2021 , पृष्ठ 23)
   - एनआरसी (1999 , पृ. 210-211)
   - न्यूक्विस्ट (1994 , पृ. 235-248)
5. ^ऊपर जायें:[ए बी] फर्स्टएआई विंटर,लाइटहिल रिपोर्ट,मैन्सफील्ड संशोधन
   - क्रेवियर (1993 , पृ. 115-117)
   - रसेल और नॉरविग (2021 , पीपी. 21-22)
   - एनआरसी (1999 , पृ. 212-213)
   - होवे (1994)
   - न्यूक्विस्ट (1994 , पृ. 189-201)
6. ^ऊपर जायें:[ए बी] दूसराएआई विंटर:
   - रसेल और नॉरविग (2021 , पृष्ठ 24)
   - मैककॉर्डक (2004 , पृ. 430-435)
   - क्रेवियर (1993 , पृ. 209-210)
   - एनआरसी (1999 , पृ. 214-216)
   - न्यूक्विस्ट (1994 , पृ. 301-318)
7. ^ऊपर जायें:[ए बी] गहन शिक्षणक्रांति,एलेक्सनेट:
   - रसेल और नॉरविग (2021 , पृष्ठ 26)
   - मैकिन्से (2018)
8. ^ऊपर जायें:[ए बी सी] कृत्रिम सामान्य बुद्धि:
   - रसेल और नॉरविग (2021 , पीपी. 32-33, 1020-1021)

आधुनिक संस्करण के लिए प्रस्ताव:

- पेनाचिन और गोएर्टज़ेल (2007)

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

प्रमुख शोधकर्ताओं की ओर से AI में अतिविशेषज्ञता की चेतावनियाँ:

- निल्सन (1995)
- मैक्कार्थी (2007)
- बील और विंस्टन (2009)

9. ^ रसेल और नॉरविग (2021 , §1.2)
10. ^ समस्या समाधान, पहेली समाधान, खेल खेलना और कटौती:
    - रसेल और नॉरविग (2021 , अध्याय 3-5)
    - रसेल और नॉरविग (2021 , अध्याय 6) (बाधा संतुष्टि)
    - पूले, मैकवर्थ और गोएबेल (1998 , अध्याय 2,3,7,9)
    - लुगर और स्टबलफ़ील्ड (2004 , अध्याय 3,4,6,8)
    - निल्सन (1998 , अध्याय 7-12)
11. ^ अनिश्चित तर्क:
    - रसेल और नॉरविग (2021 , अध्याय 12-18)
    - पूले, मैकवर्थ और गोएबेल (1998 , पृ. 345-395)
    - लुगर और स्टबलफ़ील्ड (2004 , पृ. 333-381)
    - निल्सन (1998 , अध्याय 7-12)
12. ^ऊपर जायें:[ए बी सी] अघुलनशीलता और दक्षताऔरदहनशील विस्फोट:
    - रसेल और नॉरविग (2021 , पृ. 21)
13. ^ऊपर जायें:[ए बी सी] उप-प्रतीकात्मक तर्क और ज्ञान की व्यापकता के मनोवैज्ञानिक साक्ष्य:
    - कन्नमैन (2011)
    - ड्रेफस और ड्रेफस (1986)
    - वासन और शापिरो (1966)
    - कन्नमैन, स्लोविक और टावर्सकी (1982)
14. ^ ज्ञान प्रतिनिधित्व और ज्ञान इंजीनियरिंग :
    - रसेल और नॉरविग (2021 , अध्याय 10)
    - पूले, मैकवर्थ और गोएबेल (1998 , पृ. 23-46, 69-81, 169-233, 235-277, 281-298, 319-345)
    - लुगर और स्टबलफ़ील्ड (2004 , पृ. 227-243),
    - निल्सन (1998 , अध्याय 17.1-17.4, 18)
15. ^ स्मोलियार और झांग (1994) ।
16. ^ न्यूमैन और मोलर (2008) ।
17. ^ कुपरमैन, रीचले और बेली (2006) ।
18. ^ मैकगैरी (2005) ।
19. ^ बर्टिनी, डेल बिम्बो और टोर्नियाई (2006) ।
20. ^ रसेल और नॉरविग (2021) , पीपी 272।
21. ^ श्रेणियों और संबंधों का प्रतिनिधित्व: सिमेंटिक नेटवर्क , विवरण तर्क , विरासत (फ्रेम और स्क्रिप्ट सहित):
    - रसेल और नॉरविग (2021 , §10.2 और 10.5),
    - पूले, मैकवर्थ और गोएबेल (1998 , पृ. 174-177),
    - लुगर और स्टबलफ़ील्ड (2004 , पृ. 248-258),
    - निल्सन (1998 , अध्याय 18.3)

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 9, September 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1109018 |

22. ^ घटनाओं और समय का प्रतिनिधित्व: स्थिति कैलकुलस , इवेंट कैलकुलस , धाराप्रवाह कैलकुलस (फ़्रेम समस्या को हल करने सहित):
    - रसेल और नॉरविग (2021 , §10.3),
    - पूले, मैकवर्थ और गोएबेल (1998 , पृ. 281-298),
    - निल्सन (1998 , अध्याय 18.2)
23. ^ कारण कलन :
    - पूले, मैकवर्थ और गोएबेल (1998 , पृ. 335-337)
24. ^ ज्ञान के बारे में ज्ञान का प्रतिनिधित्व: विश्वास कैलकुलस, मोडल लॉजिक्स :
    - रसेल और नॉरविग (2021 , §10.4),
    - पूले, मैकवर्थ और गोएबेल (1998 , पृ. 275-277)
25. ^ऊपर जायें:[ए बी] डिफ़ॉल्ट तर्क,फ़्रेम समस्या,डिफ़ॉल्ट तर्क,गैर-मोनोटोनिक तर्क,परिधि,बंद दुनिया धारणा,अपहरण:
    - रसेल और नॉरविग (2021 , §10.6)
    - पूले, मैकवर्थ और गोएबेल (1998 , पृ. 248-256, 323-335)
    - लुगर और स्टबलफ़ील्ड (2004 , पृ. 335-363)
    - निल्सन (1998 , ~18.3.3)

    (पूले और अन्य अपहरण को "डिफ़ॉल्ट तर्क" के अंतर्गत रखते हैं। लुगर और अन्य इसे "अनिश्चित तर्क" के अंतर्गत रखते हैं)।

26. ^ऊपर जायें:[ए बी] सामान्य ज्ञान ज्ञान का विस्तार:
    - लेनात और गुहा (1989 , परिचय)
    - क्रेवियर (1993 , पृ. 113-114),
    - मोरावेक (1988 , पृष्ठ 13),
    - रसेल और नॉरविग (2021 , पीपी. 241, 385, 982) (योग्यता समस्या)
27. ^ न्यूक्विस्ट (1994) , पृ. 296.
28. ^ क्रेवियर (1993) , पीपी 204-208।
29. ^ गर्टनर (2023) ।
30. ^ रसेल और नॉरविग (2021) , पृ. 528.
31. ^ स्वचालित योजना :
    - रसेल और नॉरविग (2021 , अध्याय 11),
32. ^ स्वचालित निर्णय लेना , निर्णय सिद्धांत :
    - रसेल और नॉरविग (2021 , अध्याय 16-18)
33. ^ शास्त्रीय योजना :
    - रसेल और नॉरविग (2021 , धारा 11.2)
34. ^ सेंसर रहित या "अनुरूप" योजना, आकस्मिक योजना, पुनर्योजना (उर्फ ऑनलाइन योजना):
    - रसेल और नॉरविग (2021 , धारा 11.5)
35. ^ अनिश्चित प्राथमिकताएँ:
    - रसेल और नॉरविग (2021 , धारा 16.7)

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details