

power of AI by defeating two human champions [14]. However, if two intelligent machines compete in playing a game (such as chess, game of cards or jeopardy), it will be interesting to observe the outcome when the environment unpredictably changes in terms of game rules, board size, etc. The author noted that in such situations, who will win may depend on which AI technique can take correct dynamic decisions (based on current state of the board, best moves and future predictions) quickly; however, *if the competing AI techniques are equally efficient, the game may result in draws.*

The Watson's success in playing games with human, led IBMers to develop an AI-powered medical technology [14] i.e. developing a general-purpose IA doctor for healthcare with the goal to reduce diagnosis errors, optimize treatment and helping doctors to identify diseases quickly. So far IBM research achieved very limited success but these are more like AI assistants that can perform certain routine tasks [15, 16]. Some AI-based tools are available in the area of medical imaging (analyzing X-rays, retina scan, tumors, etc.). For example, FDA approved [17] a device (IDx-DR) uses an AI algorithm to analyze retinal images to determine if the patient has diabetic retinopathy or not. If a positive result is detected, patients are recommended see an eye care provider for further diagnostic evaluation and possible treatment as soon as possible. Similarly, oncologists recently applied AI-based technique to analyze genetic tests for finding patterns in certain tumors (showing a deficiency in homologous DNA repair), which makes them susceptible to treatment by PARP inhibitors [18]. The author argues that patients' medical (observable) data are noisy, incomplete, and their distribution in feature space may not be separable for an AI/ML to learn patterns from historical medical records to build an autonomous AI doctor for accurate classification/diagnosis. Until medical field understand all the causes of a disease (that are measurable) and can be mapped properly to AI algorithms, human experience (in inferring from different test results, clinical reports, physical checkups, and experiences) produce, in general, better diagnosis i.e. using doctor-in-the-loop.

Many AI tools are available (such as *weka*, *tpot*, *tensorflow*, *alexnet*, *pytorch*, *keras*, *scikit*, etc.) for practitioners' use. So AI applications continue to spread in many domains ranging from natural language processing to voice recognition, sentiment analysis, drug discovery, event correlation, market and climate change predictions, business analytics, etc. Companies of all sizes including Google and Facebook are harnessing AI for developing products and services, and claiming benefits in terms of better customer experiences and increased revenues. Business experts' argue that AI/ML in one form or another will bring fourth industrial revolution, collecting wide range of information and providing business opportunities. Many conferences and events are now being hosted to exhibit the power of AI and deep learning in industries like agriculture, logistics, retail, healthcare and more [19].

Reproducibility issues: While some of these AI tools including Deep Neural Networks (DNN) have been successfully used in pattern/image recognition and machine vision [4-8] but used as a "blackbox" which raised questions in scientific community. Dr. Genevera Allen [9, 10] recently pointed out "Growing amount of scientific research involves using machine learning software to analyse data that has already been collected. This happens across many subject areas ranging from biomedical research to astronomy producing results that are misleading, often completely wrong and causing reproducibility crisis in science." Following her comments on reproducibility, data science researchers conducted survey and reported in a blog posting [11], also a leading AI conference (NeurIPS 2019) introduced reproducibility checklist policy for publications to alleviate this crisis. The author argues that the reasons for not being able to reproduce results are because of the lack of experimental details provided, and in many publications, biased statistical results are reported which are hard to verify. In general, performances of AI algorithms largely depend on tuning various control parameters (both internal and external), mapping functions, encoding schemes, distance measures, recognition thresholds, metaheuristics, etc. and should be reported along with

performance metrics for useful reproduction. Also for a specific application, hybridization of AI techniques, data pre-processing (sampling and dimensionality reduction) and post-processing (output filtering, and visual interpretation) play important roles in algorithmic success, and in many cases lack to include in reporting.

Dual-role of AI: Like many other technologies, *AI can play dual-role* and can be used in many different ways with varying intent. For example:

- AI-based robots/agents are becoming integral part of human body through implants and precision surgery; these tools are also generating lucrative business for entertainment, pleasure and the workforce. Interestingly, such technology uses for performing physical activities while making significant changes to man-machine connections but if abused can also cause human disaster.
- Online search algorithms (software agents [20]) are being used for finding best offers/discounts for hotels, air tickets, goods and services from different vendors/sellers, as long as such searches performed to make decisions are independent of sharing information with the competitors, the process seems ethical and legal. On the other hand, collaborative multi-agent systems can autonomously coordinate with sellers for market price fixing or gouging and establish trade terms (under some market conditions) without human involvement aiding anti-competitiveness. Laws prohibits a company and its competitors engage in cartel to set dynamic pricing and the use AI agents is not immune to such a practice [21].
- AI is being used for composing essays and stories, and checking answer papers. In particular, AI-based multi-talented, natural language processing algorithms called BERT and Elmo were developed to understand and answer questions [22-24]. Similarly, the OpenAI research group has demonstrated an AI-based software that can compose authentic-looking fake news articles from a few pieces of information about the intended story [25].
- AI-based image morphing apps (Face2Face, FaceApp) can automatically modify someone's face to add a smile, make younger or older looking faces, or swap genders. Such apps can provide "beautifying" effects that include smoothing out wrinkles and, more controversially, lightening the skin. AI tools are also making it easy to generate realistic videos, and impersonate someone so that a person's facial expressions match those of someone being tracked using a depth-sensing camera [26]. Some AI tools (such as Lyrebird) can be used to impersonate another person's voice and manipulate reality.
- While AI methods are deployed to detect spread of misinformation, AI can also be used to covertly increase the number of likes and downloads of fake news posting in social media, such unethical use of AI resulting in diminishing trust of open source information on the Internet [27].

Defensive AI: AI techniques [28] are used for cyber security and privacy, in particular, for situation awareness/monitoring such as intrusion/anomaly detection and response, virus/malware detection, spam/spyware/adware detection and filtering, adaptive firewalls, attack prevention or repair, etc.

Offensive AI: When AIs are used for malicious purposes [29-30] including but not limited to obfuscate, hiding communications, evolving malware, finding security holes in software and systems through reconnaissance, releasing specialized bots, launching targeted attacks, data breach, stealing personally-identifiable information (PII), etc.



Figure 2: Illustration of offensive and defensive AI in cyber warfare.

The figure 2 illustrates that if two AI-based systems or (ro)bots compete each other to win a race, game, fight, or takeover, who will win depends on their algorithmic complexity and flexible strategies as incorporated by the developers (human intelligence). It is to be noted that all that can be done with AI to defend cyber/physical systems, the same can be achieved to defeat them. For example, while an AI can be used to detect spam, another AI can be used to defeat that spam detector (exploiting the detection patterns). Similarly, AI-based situation monitoring systems are in use for surveillance; such systems can also be used for sniffing and spying with some parametric changes.

According to the Security Expert, Bruce Schneier [31], “Finding software vulnerabilities can benefit both attackers and defenders, but it's not a fair fight. When an attacker's ML system finds a vulnerability in software, the attacker can use it to compromise systems. When a defender's ML system finds the same vulnerability, he or she can try to patch the system or program network defenses to watch for and block code that tries to exploit it.” AI-Based Fuzz testing are in use for a decade to identify bugs in software and web applications, where random or unexpected inputs are generated in an attempt to cause exceptions, crash or halt the service. The purpose of such testing is to identify inputs which may cause undefined/erroneous behavior and fix these before release; hackers also use variety of Fuzzers to exploit them upon release. Some bug bounty frameworks use AI approaches to test code coverage capabilities and finding as many errors and identify the paths in the target software.

Adversarial Machine Learning (AML) [32] and Generative Adversarial Networks (GAN) [33] are being used to train supervised learning algorithms (such as Deep Neural Networks) in order to build robust decision support systems. Interestingly, different variants of these tricks can also use by hackers to defeat such data-driven trained AI systems. Specifically, AI-based techniques can even play multiple primary and secondary roles with conflicting objectives, where the purpose may be to create disruption or chaos; one such example is the use of social media intelligent bots for trolling.

AI Challenges: As mentioned earlier, most AI/ML techniques are data-driven and rely on data quality; these are used to analyze historic/current data and behaviors and predict events/trends to decide an appropriate response, they can also detect deviations/anomalies from past or forecasted behaviors. For instance, playing games (either by two parties or multi-party collaborative game) by sharing incomplete/misinformation for deception, where the use of AI logics become challenging [34]. Similarly, in cyber deception, the system defender neither can directly determine whether a user is legitimate nor can observe the user’s malicious behavior in multi-stage deception games [35]. Moreover, if the environment is unpredictably dynamic, uncertain, misleading, and have man-made obfuscation where behavior profiling or knowledge patterns are difficult to harness, most AI/ML techniques in practice today will miserably fail. For example, autonomous self-controlled weapons can occasionally trigger or misfire in unpredictable ways

causing collateral damages. Human Rights Watch's Mary Wareham [36] “denounced AI systems for warfare autonomous weapons because machines are not moral agents and so cannot be responsible for making decisions of life and death”. For example, military aircraft and drone that take off, fly, and land on their own and robotic sentries that can identify movement are limited by AI's abilities in image recognition and do not have the detail or context to be judge, jury, and executioner on a battlefield [36].

Future of AI: Though most AI/ML (brain models) are nature and bio-inspired techniques and exist for more than three decades, their growing success in commercialization come mainly due to today's computing power, storage capacities, virtualization, bigdata platforms, network bandwidth, scripting languages, user-friendly application (App) frameworks and wide variety of cloud services. All devices, tools and automated services are becoming smarter whether these are home appliances, robots, vehicles, network gears, 5G/6G communication or other applications; intelligent techniques will continue to play very important roles. It is to be noted that AI algorithms rely on mathematical logics and associated control parameters; studies show that combining heuristics and hybridization of AI techniques, work better in fine-tuning solutions to complex real-world problems. With the progress of Explainable AI (XAI) research, it will be possible to provide interpretability on algorithmic biasness, reason for producing a decision, etc. and (not to treat an AI tool as a “black box”) can easily be understood by applicators and domain experts [37]. However, further research will be needed to develop advanced AI techniques which can provide accurate decisions under a wide variety of unexpected, uncertain and hostile situations. New research should focus on open world problems having varying degrees and types of uncertainties (at the local and global levels) that violate implicit or explicit assumptions of the environment. The author points out that there are ample evidences in natural and biological systems and processes on how to handle the imminent danger and/or survive in unpredictable situations which can provide guidance to design innovative AI strategies [38]. The critics may argue that these processes are relatively slow and not applicable to AI technologies but if we can better understand these natural/biological processes and mechanisms, developing and implementing some of these strategies in a faster computing environment will not be difficult and will alleviate the solutions. Recently, new research program launched by NSF [39], called Real-Time Machine Learning (RTML) for developing novel hardware-software (ML) architecture which can learn from continues stream of new data and make decisions in real-time as incremental, one-shot or online learning systems.

AI for greater good: The use of AI-based approaches for intelligence gathering and counter-intelligence are on the raise; the world will witness cyber/AI arm-race which will have adverse effect on trust [40]; particularly, it will affect knowledge globalization, software/hardware supply chain and collaborative innovations. Similar to international nuclear or bio/chemical agent treaties, use of AI-based weapons, harmful (ro)bots, DNA manipulation and cyber-attacks, treaties be needed to prevent development and proliferation of AI-based malicious activities and destructions. Several reports highlighted the potential security threats from malicious uses of AI technologies [41, 42]. The US lawmakers introduced a bill, called the *Algorithmic Accountability Act* [43] that would require companies to audit their AI/ML systems for bias and discrimination in order to regulate the bad use of AIs. The European Union has crafted *seven principles* for guiding AI development and building trust. While these guidelines are not binding, these will provide the basis for further actions [44]. Mariya Gabriel, Europe's top official on the digital economy, said to companies using AI "People need to be informed when they are in contact with an algorithm and not another human being. Any decision made by an algorithm must be verifiable and explained." We hope that XAI research can provide some light in that effort.

In closing, it is predicted that the business adoption of AI will increase significantly in next two years, and International Data Corp. projected that global AI spending will reach \$79.2 billion by 2022. Considering unparalleled benefits of AIs and their possible misuse/abuse, regulations become essential for AI-based developers to take responsibilities of their products. Accordingly, proper guidelines be needed for best practices and ethical principles to prevent the development of AI-based programs and systems prone to be

misused or bad for the humanity. In December 2016, IEEE has taken global Initiative for ethically aligned design of Artificial Intelligence tools and Autonomous Systems (AI/AS) that encourages technologists to prioritize ethical considerations of Human Wellbeing; more than hundred global thought leaders and experts of artificial intelligence, ethics, and related fields joined this effort [45].

References:

1. Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Book). First published in 1995 and the third edition of the book was released on December 2009.
2. Martin Hilbert. *Big Data for Development: A Review of Promises and Challenges*. Development Policy Review. Wiley Online Library, December 2015.
3. M Chen, S Mao, Y Liu. *Big data: A survey*, Mobile networks and applications, Springer, 2014.
4. James Glass and Yonaton Belinkov. Putting neural networks under the microscope. Reported by Rob Matheson, MIT News Office, February 1, 2019.
5. Josh Patterson and Adam Gibson. *Deep Learning: A practitioner's Approach*. O'Reilly Press, 2017. ISBN: 978-1-491-91425-0
6. David Harwath, Adria Recasens, Didac Suris, Galen Chuang, Antonio Torralba, and James Glass. Jointly Discovering Visual Objects and Spoken Words from Raw Sensory Input. Published in the proceedings of European Conference on Computer Vision (ECCV 2018) and MIT Press, September 2018. <http://news.mit.edu/machine-learning-image-object-recognition-0918>.
7. AI-powered horse monitoring: Magic AI raises \$1.2M to expand equine technology. Magic AI Corp: <https://www.magicai.io/>
8. F. J. de Cos, M Woźniak, J. A. Méndez, J. R. V. Flecha. Special issue SOCO 2017: AI and ML applied to Health Sciences (MLHS). In *Journal of Neural Computing and Applications*- Springer, 2019 <https://link.springer.com/article/10.1007/s00521-019-04031-0>.
9. Genevera Allen, AAAS: Machine learning 'causing science crisis'. Reported by Pallab Ghosh, Science correspondent, BBC News, Washington, posted on 16 February, 2019. <https://www.bbc.com/news/science-environment-47267081>.
10. Genevera Allen. Can we trust scientific discoveries made using machine learning? Rice University Press Release on February 18, 2019, <http://news.rice.edu/2019/02/18/can-we-trust-scientific-discoveries-made-using-machine-learning/>
11. Blog Posting: A quick response to Genevera Allen about Machine learning 'causing science crisis' <https://towardsdatascience.com/a-quick-response-to-genevera-allen-about-machine-learning-causing-science-crisis-8465bbf9da82>.
12. Murray Campbell, A. Joseph Hoane Jr., Feng-hsiung Hsu. Deep Blue. *Artificial Intelligence*, 134, PP. 57–83, 2002.
13. 20 Years after Deep Blue: How AI Has Advanced Since Conquering Chess. <https://www.scientificamerican.com/article/20-years-after-deep-blue-how-ai-has-advanced-since-conquering-chess/>
14. IBM to Collaborate with Nuance to Apply IBM's "Watson" Analytics Technology to Healthcare, February 2011 (<https://www-03.ibm.com/press/us/en/pressrelease/33726.wss>).
15. Heal Thyself. A New report on IBM Watson, *IEEE Spectrum*, April 2019.
16. Robert Wachter. *The Digital Doctor: Hope, Hype, and Harm at the Dawn of Medicine's Computer Age*, Book, April 1, 2015.
17. FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems, FDA News Release, April 11, 2018. <https://www.fda.gov/newsevents/newsroom/pressannouncements/ucm604357.htm>
18. AI Test Pinpoints More Cancers Targeted by Astra, Glaxo Drugs. By John Lauerman April 15, 2019, *Bloomberg News*, <https://www.bloomberg.com/news/articles/2019-04-15/ai-test-pinpoints-more-cancers-targeted-by-astra-glaxo-drugs>

19. Evolving Industries with Intelligent Machines and IoT. https://www.nvidia.com/en-us/gtc/topics/iot-and-intelligent-machines/?ncid=pa-pai-gcsj28-32982&gclid=EAIaIQobChMIiNT3g6yo4AIVkbrACh1UKwM-EAAYAAEgIGVfD_BwE.
20. Jeffrey O. Kephart, James E. Hanson, Amy R. Greenwald. Dynamic pricing by software agents. In *Journal of Computer Networks*, Elsevier publication, 32, 731-752, 2000.
21. Pricing algorithms can learn to collude with each other to raise prices. MIT Technology Review (Reporter Karen Hao), February 2019. <https://www.technologyreview.com/f/612947/pricing-algorithms-can-learn-to-collude-with-each-other-to-raise-prices/>
22. Matthew E. Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, Luke Zettlemoyer. Deep contextualized word representations, In *North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT)*, New Orleans, Louisiana, June 1-6, 2018.
23. C Alberti, K Lee, M Collins. A BERT baseline for the natural questions. In *Computation and Language (cs. CL)*, arXiv preprint arXiv:1901.08634, 2019 - arxiv.org.
24. J Liu, Y Xu, L Zhao. Automated Essay Scoring based on Two-Stage Learning-preprint arXiv:1901.07744, 2019 - arxiv.org
25. The AI That Can Write a Fake News Story From a Handful of Words. By Jeremy Kahn, February 14, 2019, <https://www.bloomberg.com/news/articles/2019-02-14/the-ai-that-can-write-a-fake-news-story-from-a-handful-of-words>.
26. Real or Fake? AI Is Making It Very Hard to Know. MIT Technology Review by Will Knight, May 1, 2017. <https://www.technologyreview.com/s/604270/real-or-fake-ai-is-making-it-very-hard-to-know/>
27. Elle Hunt. What is fake news? How to spot it and what you can do to stop it. *The Guardian*. December 17, 2016. <https://www.theguardian.com/media/2016/dec/18/what-is-fake-news-pizzagate>
28. D. Dasgupta, Program Chair, IEEE Symposium on Computational Intelligence in Cyber Security at SSCI, 2007-2019.
29. Miles Brundage (and 25 other co-authors).The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation, February 2018, e-print achieved at arxiv.org.
30. Federico Pistono, Roman V. Yampolskiy. Unethical Research: How to Create a Malevolent Artificial Intelligence, in *Proceeding of AAAI 2018*, e-Print archived at arxiv.org
31. Bruce Schneier. Machine Learning to Detect Software Vulnerabilities. *Schneier on Security*, 2019. https://www.schneier.com/blog/archives/2019/01/machine_learnin.html.
32. L Huang, A D Joseph, B Nelson, B. I. P. Rubinstein and J.D. Tygar. Adversarial machine learning. In *Proceedings of the 4th ACM workshop on Security and artificial intelligence*, Pages 43-58, Chicago, Illinois, USA, October 21 - 21, 2011
33. I Goodfellow, J Pouget-Abadie, M Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative adversarial nets. In *Advances on Neural Information Processing Systems (NIPS)*, issue 27, 2014.
34. D. Dasgupta, P. Vejjandla, A. Kaushal, F. Nino. Evolving attacker-defender gaming strategies in a simulated network. In the proceedings of the Second IEEE International Conference on Information Privacy, Security, Risk and Trust (PASSAT), Minneapolis, Minnesota, USA, August 20-22, 2010.
35. Tao Zhang, Linan Huang, Jeffrey Pawlick, and Quanyan Zhu. Game-Theoretic Analysis of Cyber Deception: Evidence Based Strategies and Dynamic Risk Mitigation, Published in ArXiv 2019
36. Call to ban killer robots in wars. By Pallab Ghosh, Science correspondent, BBC News, February 15, 2019. Washington DC, <https://www.bbc.com/news/science-environment-47259889>.
37. Filip Dositovic, Mario Brcic and Nikica Hlupic. "Explainable Artificial Intelligence: A Survey". In proceedings of 41st International Convention on MIPRO, Croatia. pp. 210–215, 2018.
38. Dipankar Dasgupta and Fernando Nino, (authors). *Immunological Computation: Theory and Applications (Book)*, CRC press, September 2008.
39. NSF Solicitation (#19-566) on Real-Time Machine Learning (RTML) program, June 2019, <https://www.nsf.gov/pubs/2019/nsf19566/nsf19566.htm>
40. Dipankar Dasgupta, Denise M. Ferebee. Consequences of Diminishing Trust in Cyberspace. In the proceeding of the 8th International Conference on Information Warfare and Security (ICIW), Denver, Colorado, March 25-26, 2013.

41. Peter J. Bentley, Miles Brundage, Olle Häggström, Thomas Metzinger. Should we fear artificial intelligence? In-depth Analysis, European Union, Scientific Foresight Unit (STOA), March 2018
42. E. Oliveira. Beneficial AI: the next battlefield. In Journal of Innovation Management, issue no.5, vol. 4, pp 6-17, 2017. ISSN 2183-0606. <https://journals.fe.up.pt/index.php/IJMAI/article/viewFile/501/307>.
43. Congress wants to protect you from biased algorithms, deepfakes, and other bad AI. MIT Technology Review (reported by Karen Hao on April 15, 2019). Originally published in *Webby-nominated AI newsletter The Algorithm*. <https://www.technologyreview.com/s/613310/congress-wants-to-protect-you-from-biased-algorithms-deepfakes-and-other-bad-ai/>
44. Europe is making AI rules now to avoid a new tech crisis. By Ivana Kottasová, CNN Business. Updated April 8, 2019. <https://www.cnn.com/2019/04/08/tech/ai-guidelines-eu/index.html>
45. IEEE Global initiative on Ethically Aligned Design. <https://standards.ieee.org/industry-connections/ec/autonomous-systems.html>.

About the Author:

Dr. Dipankar Dasgupta is a professor of Computer Science at the University of Memphis since 1997, an IEEE Fellow and an ACM Distinguished Speaker. Dr. Dasgupta is known for his pioneering work on the design and development of intelligent solutions inspired by natural and biological processes. During 1990-2000, he extensively studied different AI/ML techniques and research in the development of an efficient search and optimization method (called structured genetic algorithm) has been applied in engineering design, neural-networks, and control systems. He is one of the founding fathers of the field of artificial immune systems (a.k.a Immunological Computation) and is at the forefront of applying bio-inspired approaches to cyber defense. His notable works in digital immunity (featured in Computer World Magazine 2003), negative authentication, cloud insurance modeling and adaptive multi-factor authentication demonstrated the effective use of various AI algorithms.



Dr. Dasgupta has authored four books, his latest graduate textbook is on Advances in User Authentication published by Springer-Verlag, August 2017. In addition, Dr. Dasgupta has published more than 260 research papers (+15,000 citations as per google scholar) in book chapters, journals, and international conference proceedings. Among many awards, he was honored with the 2014 ACM-SIGEVO Impact Award for his seminal work on negative authentication, an AI-based approach. He also received five best paper awards in different International conferences and has been organizing IEEE Symposium on Computational Intelligence in Cyber Security at SSCI since 2007. Dr. Dasgupta is an ACM Distinguished Speaker, regularly serves as panelist and keynote speaker and offer tutorials in leading computer science conferences, and have given more than 300 invited talks in different universities and industries.