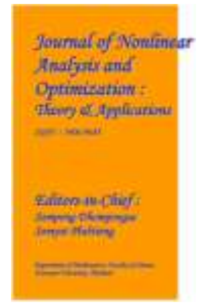
Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 14, Issue. 01 : 2023
ISSN : **1906-9685**



EXPLAINABLE AI AND ITS IMPORTANCE IN DECISION MAKING

Himanshu Arora

Professor

Computer Science Engineering

Arya Institute of Engineering & Technology, Jaipur, Rajasthan

Minika Lamba

Associate Professor

Electronics & Communication Engineering

Arya Institute of Engineering & Technology, Jaipur, Rajasthan

Lokesh kumar
Science Student

Vivekanand Senior Secondary School, Rewari, Haryana

Sameer Kumar
Science Student

Lord Buddha Public School, Aurangabad, Bihar

http://doi.org/10.36893/JNAO.2023.V14I1.0019-0027

Abstract

AI is the technology that builds intelligent machines able to perform tasks that generally need human intelligence in machines that are programmed to think and act like humans. AI has penetrated many organization processes, resulting in a growing fear that intelligent machines will soon replace many humans in decision-making. To provide a more proactive and pragmatic perspective, this article highlights the complementarities of humans and AI. It examines how each can strengthen organizational decision-making processes typically characterized by uncertain, complex and equal vocalists. With excellent computation information processing capacity and an analytical approach, AI can extend human cognition when addressing complexity. In contrast, humans can still offer a more holistic, intuitive approach in dealing with uncertainly and equal vocalist in organizational decision-making.

Keywords

Artificial intelligence, decision making, human machine, equal vocalist

I. Introduction

Technologies related to artificial intelligence (AI), particularly machine learning (ML), are becoming more and more commonplace and are being applied to a variety of jobs. Artificial Intelligence (AI) technologies, particularly Machine Learning (ML), are becoming more and more commonplace despite the fact that algorithms can perform remarkable performance, full delegation to ML models is sometimes not sought because of their probabilistic character, and nothing can be guaranteed. This data may contain biases, unidentified defects, and input errors. Machine learning models can help human decision-makers create cooperation. There are numerous obstacles to overcome while using AI.

The power dynamics among governments, corporations, and citizens can be significantly impacted by the decisions made by autonomous computational. It is getting more difficult to recognize and understand algorithms as they get more independent and opaque. The behavior of an algorithmic society may be too opaque for it to be held responsible for. The public should be informed about the decisions made by computer algorithms, what information they consider, and how they arrive at these conclusions.

http://doi.org/10.36893/JNAO.2023.V14I1.0019-0027

AI lacks a widely agreed-upon definition. It is commonly known as a machine's capacity to learn from mistakes, adapt to novel inputs, and carry out tasks that resemble those of a human. In the 1950s, the terms artificial intelligence (AI) and AI systems were first used. AI has since gone through ups and downs, or "AI springs" and "AI winters." The availability and power of Big Data are revitalizing artificial intelligence (AI), thanks to big data's rapid growth in technologies such as enhanced computer storage capacity and ultra-fast data processing equipment.

With their remarkable capacity for self-improvement, emerging AI systems such as Watson are becoming more and more useful for knowledge-based jobs that were previously thought to be the sole domain of humans. Previously, these jobs were thought to be immune to automation because they were carried out by white-collar individuals. Artificial intelligence (AI) systems are becoming more intelligent at a rapid pace and are making decisions in a wider range of difficult scenarios with a degree of autonomy. Sophisticated smart technologies are poised to supplant human labor in a multitude of domains, ushering in a "second machine age" for postindustrial economies. With their remarkable capacity for self-improvement, emerging AI systems such as Watson are becoming more and more useful for knowledge-based jobs that were previously thought to be the sole domain of humans. Previously, these jobs were thought to be immune to automation because they were carried out by white-collar individuals. Artificial intelligence (AI) systems are becoming more intelligent at a rapid pace and are making decisions in a wider range of difficult scenarios with a degree of autonomy (Davenport & Kirby, 2016). Sophisticated smart technologies are poised to supplant human labor in a multitude of domains, ushering in a "second machine age" for postindustrial economies.

II. Previous Work/Literature Review-

The foundation of expert systems is a vast set of rules designed to approximate the knowledge of an expert. Thus, the term "expert system" is employed. Usually, rules are expressed as implications, which allow for the derivation of new conclusions in the event that certain premises are true. A trace of the application of rules with consistent premises and conclusions makes up an explanation. According to one explanation, "the system concluded that the patient has this illness because it applied these rules in this order to these initial symptoms." These are examples of trace explanations. The fundamental characteristic shared by all of the aforementioned systems is their foundation in symbolic representation. Symbolic systems use human-understandable languages (symbols) that humans may use to check reasoning even if they are designed for machine processing. When making decisions, the logic used to reach a conclusion simulates human reasoning, for example, by using the "if... then..." principles. These regulations frequently take the shape of human-written computer code. Moreover, the logic of symbolic systems may be easily understood by specialists, but not by non-experts. As a result, XAI frequently focuses on determining whether the results are accurate. Therefore, occasionally interpretable machine learning is used instead of XAI.

TABLE I: - Terms used in the full-text search of the SSIS/IJIM archive

case-based reasoning		11.logic programming
computer vision		12.machine learning
cognitive computing		13.machine vision
cognitive science		14.natural language processing
data mining		15.neural network
data science		16.pattern recognizing
expert system		17.recommendation system
fuzzy logic		18.semantic network
genetic algorithm		20.speed recognizing
k-means		

Papers where the phrase "intelligent" had no relation to artificial intelligence (AI) and those where the term had only been identified in a reference that was quoted without mentioning AI were excluded from further screening. This resulted in 123 SSIS/IJIM papers overall. We divided them into two categories: those where AI was discussed in-depth in 52 of the papers, and those where it was only addressed briefly in 71. The papers in the latter category ranged from those that mentioned artificial intelligence (AI) as merely one sort of system in a list of systems that an organization uses or might use. Even the articles in the second category, in our opinion, merited being included in the tallies.

Our goal was to establish a situation that resembles actual AI-assisted decision-making situations, when people possess similar domain expertise to the AI and are driven to maximize the decision's outcome. Two steps were taken to strengthen the ecological validity.

http://doi.org/10.36893/JNAO.2023.V14I1.0019-0027

First, in addition to a base pay of $3, the decision performance was tied to a monetary bonus, with a reward of 5 cents if the final prediction was accurate and a loss of 2 cents if not. Previous studies have demonstrated the effectiveness of such a reward strategy in encouraging participants to maximize the choice outcome.

Second, we improved the domain knowledge and performance of MTurk workers—who were probably unfamiliar with this task—by assigning a training task and providing them with an extra piece of information. This information is the third column in Figure 1, which indicates the likelihood that an individual with that attribute-value will earn more than $50,000 on a scale of 0 to 10. Using data from the training dataset, this chance number was computed by dividing the total income of individuals with the matching attribute-value by 50K. Because previous research indicates that individuals comprehend frequencies better than probabilities, we rounded the numbers and multiplied the percentages by 10.

It is crucial to note, nevertheless, that the attribution of bias in this instance assumes that the decision model's objective is to maximize the distribution of medical resources against a baseline probability of mortality that is unrelated to the state of medicine today. However, this is probably not a reasonable assumption for many patients, not only asthmatics, to the extent that the training data reflect the chance of death given current medical management. The results of other practices that are not included in the statistics are reflected in each person's outcomes. However, patients are likely to receive varying degrees of care depending on their co morbidities or medical history. Lacking a thorough and detailed depiction of patient characteristics and treatment methods.

This example also highlights the risks associated with mistaking the plausibility of connections in interpretable systems for intervention-exploitable causal linkages. High-precision diagnoses and forecasts can be made by machine learning algorithms by utilizing associations found in data sets.

It is erroneous to assume that those associations will follow causal linkages in a way that we can take advantage of through intervention, unless there are systems in place that are expressly tailored for causal discovery. Thus, interpretability could encourage the mistaken belief that knowing a set of connections important for particular diagnostic or prediction tasks will improve our performance on other tasks for which those associations are not as useful.

III. Conclusion-

http://doi.org/10.36893/JNAO.2023.V14I1.0019-0027

1) Switch percentage or the proportion of trials where the user chose to base their final guess on the AI's prediction. When the AI's prediction was displayed, it represented the proportion of trials in which participants and the AI disagreed that used the AI's guess. It was the proportion of trials when participants opted to assign the forecast to the AI out of all trials when the AI's prediction was not displayed. Each technique can be investigated further and put to the test in real-world scenarios. Policymakers can use the strategies developed here to establish legitimacy and trust in XAI, but they can also serve as a basis for future XAI research. Because it is exploratory in nature, this research has a number of limitations. Three case studies are used to illustrate the work. Other case studies may present different difficulties. The tactics may not be all-inclusive and may be expanded upon in subsequent studies. Furthermore, neither the strategies' applicability nor our investigation into how to use them to achieve successful results was tested in real-world scenarios. The goal of XAI research should go beyond simply unlocking an algorithmic black box. A new human-machine symbiosis is required as a result of the rise of AI, which indicates a changing labor divide between humans and machines. Many conceptions of human-machine collaboration imply that humans should concentrate on more creative work while machines handle routine chores. This article goes beyond this straightforward vision and advances the idea of human-machine collaboration by highlighting the relative advantages that humans and machines hold in relation to the three characteristics that plague almost all organizational decision-making situations. This is made possible by the significant advancements in AI capabilities in recent years. While AI's capabilities aid humans in overcoming complexity through the better analytical approach of the machines, the function of human decision makers, and their ability to use intuition when handling ambiguity, particularly equivocality.

IV. Future Scope-

1. This can lead to increased efficiency and productivity in the workplace but also raises concerns about job security.

2. With advances in machine learning, robotics, and other AI technologies, more and more tasks previously performed by humans are becoming automated.

3. This trend is particularly evident in manufacturing, transportation and logistics, where robots and AI systems perform tasks like assembly, inventory management and delivery.

http://doi.org/10.36893/JNAO.2023.V14I1.0019-0027

4. AI powered automation is also impacting the job market in other ways. Some jobs being replaced by automation, while others are being transformed, with new skills and knowledge becoming essential for success in the workplace

5. One of the massive impacts of ai in our careers is job automation. As AI systems become more advanced, they can take over routine tasks and processes, leaving humans to focus on more complex tasks.

References-

[1] Amina Adadi and Mohammed Berrada. 2018. Peeking inside the black box: A survey on Explainable Artificial Intelligence (XAI). IEEE Access 6 (2018), 52138– 52160.

[2] Gagan Bansal, Besmira Nushi, Ece Kamar, Dan Weld, Walter Lasecki, and Eric Horvitz. 2019. Updates in Human-AI Teams: Understanding and Addressing the Performance/Compatibility Tradeoff. In AAAI Conference on Artificial Intelligence. AAAI.

[3] Carrie J. Cai, Jonas Jongejan, and Jess Holbrook. 2019. The effects of examplebased explanations in a machine learning interface. In International Conference on Intelligent User Interfaces, Proceedings IUI, Vol. Part F147615. Association for Computing Machinery, 258–262.

[4] Diogo V. Carvalho, Eduardo M. Pereira, and Jaime S. Cardoso. 2019. Machine Learning Interpretability: A Survey on Methods and Metrics. Electronics 8, 8 (jul 2019), 832.

[5] Hao Fei Cheng, Ruotong Wang, Zheng Zhang, Fiona O'Connell, Terrance Gray, F. Maxwell Harper, and Haiyi Zhu. 2019. Explaining decision-making algorithms through UI: Strategies to help non-expert stakeholders. In Conference on Human Factors in Computing Systems - Proceedings. Association for Computing Machinery.

[6] M. L. Cummings. 2004. Automation bias in intelligent time critical decision support systems. In Collection of Technical Papers - AIAA 1st Intelligent Systems Technical Conference, Vol. 2. 557–562.

[7] Berkeley J Dietvorst, Joseph P Simmons, and Cade Massey. 2015. Algorithm aversion: People erroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General 144, 1 (2015), 114.

[8] Jonathan Dodge, Q. Vera Liao, Yunfeng Zhang, Rachel K.E. Bellamy, and Casey Dugan. 2019. Explaining models: An empirical study of how explanations impact fairness judgment. In

http://doi.org/10.36893/JNAO.2023.V14I1.0019-0027

International Conference on Intelligent User Interfaces, Proceedings IUI, Vol. Part F147615. Association for Computing Machinery, 275–285.

[9] Finale Doshi-Velez and Been Kim. 2017. Towards A Rigorous Science of Interpretable Machine Learning. (feb 2017).

[10] Dheeru Dua and Casey Graff. 2017. UCI Machine Learning Repository.

[11] Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, and Dino Pedreschi. 2018. A survey of methods for explaining black box models. Comput. Surveys 51, 5 (aug 2018).

[12] D Harrison McKnight, Larry L Cummings, and Norman L Chervany. 1998. The Academy of Management Review. Technical Report.

[13] Tove Helldin, Göran Falkman, Maria Riveiro, and Staffan Davidsson. 2013. Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. In Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2013. .

[14] Robert R Hoffman, Shane T Mueller, Gary Klein, and Jordan Litman. 2018. Metrics for explainable AI: Challenges and prospects (2018).

[15] Jason D Johnson, Julian Sanchez, Arthur D Fisk, and Wendy A Rogers. 2004. Type of automation failure: The effects on trust and reliance in automation. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 48. SAGE Publications Sage CA: Los Angeles, CA, 2163–2167.

[16] Sharma, Richa and Kumar, Gireesh. "Availability Modelling of Cluster-Based System with Software Aging and Optional Rejuvenation Policy" Cybernetics and Information Technologies, vol.19, no.4, 2019, pp.90-100. https://doi.org/10.2478/cait-2019-0038

[17] G. Kumar and R. Sharma, "Analysis of software reliability growth model under two types of fault and warranty cost," 2017 2nd International Conference on System Reliability and Safety (ICSRS), Milan, Italy, 2017, pp. 465-468, doi: 10.1109/ICSRS.2017.8272866.

[18] Kumar, G., Kaushik, M. and Purohit, R. (2018) “Reliability analysis of software with three types of errors and imperfect debugging using Markov model,” International journal of computer applications in technology, 58(3), p. 241. doi: 10.1504/ijcat.2018.095763.

[19] Sharma, R. and Kumar, G. (2017) “Availability improvement for the successive K-out-of-N machining system using standby with multiple working vacations,” International journal of reliability and safety, 11(3/4), p. 256. doi: 10.1504/ijrs.2017.089710.

[20] Akash Rawat, Rajkumar Kaushik and Arpita Tiwari, "An Overview Of MIMO OFDM System For Wireless Communication", *International Journal of Technical Research & Science*, vol. VI, no. X, pp. 1-4, October 2021.

[21] T. Manglani, A. Vaishnav, A. S. Solanki and R. Kaushik, "Smart Agriculture Monitoring System Using Internet of Things (IoT)," *2022 International Conference on Electronics and Renewable Systems (ICEARS)*, Tuticorin, India, 2022, pp. 501-505.

[22] R. Kaushik, O. P. Mahela and P. K. Bhatt, "Power Quality Estimation and Event Detection in a Distribution System in the Presence of Renewable Energy" in Artificial Intelligence-Based Energy Management Systems for Smart Microgrids, Publisher CRC Press, pp. 323-342, 2022, ISBN 9781003290346.