Behavioral Aspects of Human-AI Collaboration in Decision-Making

Aslı Gürler-Kandemir,
Supervised by: Ayşe Kocabıyıkoğlu
Sabancı Business School

Motivation

- **Key focus:** The rise of data-driven technologies and artificial intelligence (AI) [1]

- **Implementation** in operational contexts leverage algorithms and machine-learning (ML) models to handle practitioners' complex problems, including decision-making.
- Algorithmic aids augmented in both **high-stakes decisions** and **mundane decision** contexts [2]
- **Successful implementation** of these technologies results in increased revenue and cost savings, and improved efficiency [3]

- **Without supervising** the algorithm's decision-making, their decisions may lead to **failures or mistakes** [3] 'human-in-the-loop'
- **'So what?':** "When it comes to implementation, the **success** of operations management tools […] relies heavily on our **understanding of human behavior**." [4]

Introduction

A relatively new stream of research in judgment and decision-making:

→ Different variations of key terms, not yet a clear distinction:

- Human-algorithm interaction
- Man-machine collaboration
- Algorithmic advice-taking
- **Aim:** Determining the conditions under which human and machine collaboration thrives.
- **Definition of AI/ML:** From a modern approach, which includes "tools for big data, deep learning, artificial intelligence, and other techniques for inductive prediction." [5]

Distinct human and AI characteristics (affect the decision quality):

Humans:
(1) Rely on cognitive flexibility,
(2)Synthesize information from various sources,
(3)Have limited capacity,
(4)Access to private information. [3]

AI Systems:
(1) Inherently rigid,
(2)Capable of processing a limited portion of information,
(3)Enormous quantitative capacity, [7]
(4)Efficiency and accuracy.

According to researchers, collaboration between humans and algorithmic tools is beneficial as it can outperform humans or algorithms under several conditions. [6]

Literature Review

Table 1. "Impact of the Machine on Human Decisions for Two Generic Settings" [7]

Medical assessment and diagnostic accuracy	Judicial ruling and conviction accuracy
• Overall diagnostic accuracy is improved	• Overall conviction accuracy is improved
• Fewer misdiagnosed sick patients	• Fewer acquitted guilty suspects
• More patients declared healthy when the disease incidence is high	• More suspects declared guilty when crime level is low
• More misdiagnosed healthy patients when the disease incidence is low	• More convicted nonguilty suspects when crime level is high
• Physician spends less cognitive effort to diagnose when the incidence is high	• Judge spends less cognitive effort to assess evidence when crime level is low
• Physician spends more cognitive effort to diagnose when the incidence is low and time is constrained	• Judge spends more cognitive effort to assess evidence when the crime level is high and time is constrained

Human-machine interaction can be formed in different ways:

Complementarity vs. Substitute: The decision-making process may either be co-produced with collaboration or adhered to one – human or AI.

An experimental condition [6] **of human-machine teaming:**

AI/ML Predictive Accuracy and Incentive Alignment

- **Highest** organizational profit: yielded by **human-machine teams** rather than human-alone and machine-alone.
- Second-highest: generated when the machine works alone.
- Third:-highest: when human DM works alone.

- **However,** this might not be the case for different conditions. Hence, it is important to explore different **teaming situations**.
- If the **outperformance** is proven, would decision-makers **adhere** to algorithm recommendations?

Individuals relying on machine recommendations: *algorithmic appreciation:* [8]

- People **adhere more to algorithmic advice** than human advice. [8]
- **Experts demonstrated a lower reliance** on algorithmic advice than laypeople. [8]
- **Critical question:** Under **what type of conditions** do collaborative decision-making efforts enhance a firm's performance outcomes?
- **Algorithmic bias** (e.g. socially biased outcomes from the algorithm, resulting in inequalities) [9]
- Algorithms for finding and **addressing the bias** that exists in firms [10]

Discussion

→What are the behavioral biases behind the human-machine collaboration in decision-making?

→Perception of algorithmic decision-making
→Balancing confidence in algorithmic decision accuracy

Role of the biases: shape human DM's early interpretations, laying the foundation for their future encounters (adoption)
→ **Anchoring bias:** people get affected by a proposed value while making estimations. [17]
→ **Confirmation bias:** interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand. [18]
→ **Framing bias:** not being able to decide rationally is the framing effect. [19]

Exploring Behavioral Biases

Algorithm aversion: Algorithm's forecast was more accurate than human forecasters, BUT DMs preferred the forecast done by a human, not by an algorithm. [11]

- Individuals may be **skeptical about relying on them**, although computational algorithms frequently outperform human judgment. [12]

The conditions of when and reasons why people exhibit algorithm aversion are known very little. [11]
→A challenge: a **tendency to override** algorithmic recommendations **incorrectly**.

Behavioral biases causing algorithm aversion:

(1) Mistrust bias
- For instance, in forecasting decisions, algorithms are considered better forecasters than humans. [12] **However,** research on judgmental systems such as forecasting shows that computers are trusted less than humans. [13]
- Humans have a lower tolerance for the algorithm rather than for humans for making the same mistake → quicker loss of trust in algorithms. [11]

(2) Negativity bias: the tendency for individuals to give greater weight to their observations of negative outcomes than to positive ones. [14]
- Learning failures: DM updated their belief disproportionately when the machine is incorrect compared to when the machine is accurate in its decision. [15]

(3) Selection bias:
- DM needs to decide for which decisions it can get algorithmic advice, it selects the lower-quality decisions. [16]

(4) Status-quo bias: "reluctance to change"
- Deciding the prices during clearance sales before decision-support-systems (DSS): DM's priority → minimize the inventory or considering pricing categories (for fewer pricing points).
- **However, DSS's** priority: maximize revenue regardless of inventory levels or pricing categories.
- **Observation:** Reluctancy to adhere to the DSS recommendations because they prefer to decide using their old heuristic. [3]

"If the DSS's algorithm is recommending optimal decisions but managers deviate from them, the potential improvement that the DSS was supposed to cause may never be realized."[3]

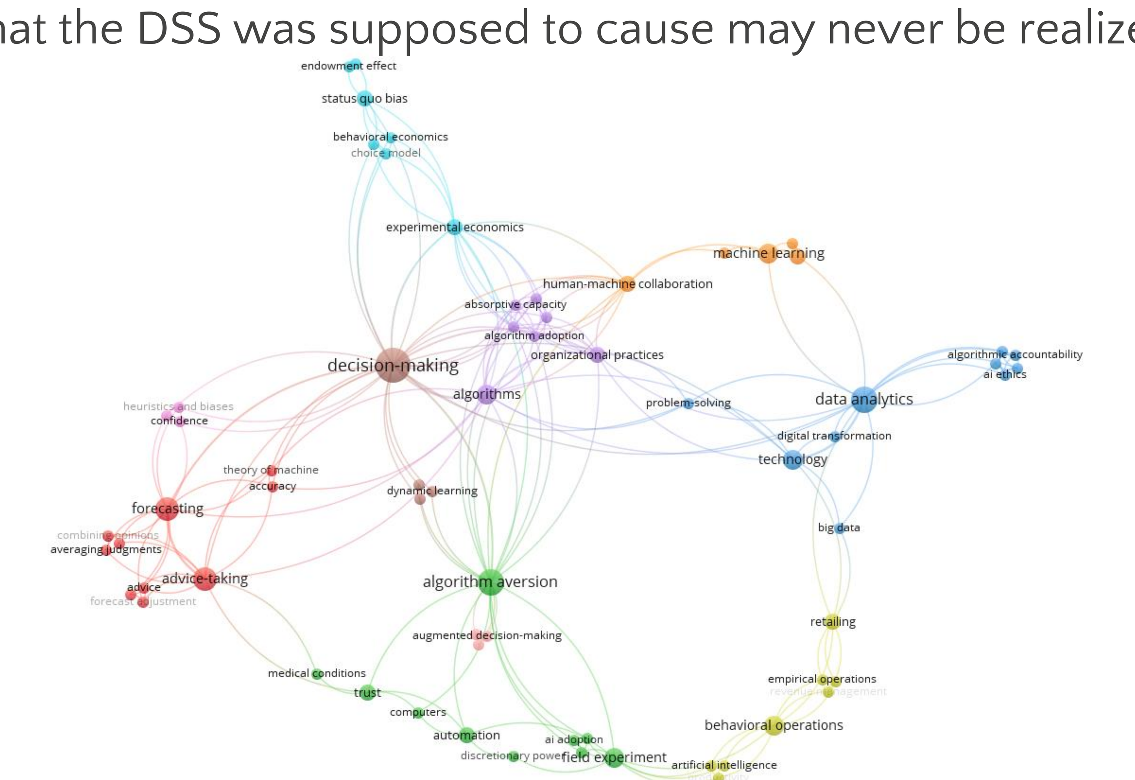

Figure 1. Visualization of keywords from the references list

References

1 Boyacı, T., Canyakmaz, C., & de Véricourt, F. (2024). Human and machine: The impact of machine input on decision making under cognitive limitations. Management Science, 70(2), 1258-1275.
2 Burton, J. W., Stein, M. K., & Jensen, T. B. (2020). A systematic review of algorithm aversion in augmented decision making. Journal of behavioral decision making, 33(2), 220-239.
3 Caro, F., & de Tejada Cuenca, A. S. (2023). Believing in analytics: Managers' adherence to price recommendations from a DSS. Manufacturing & Service Operations Management, 25(2), 524-542.
4 Bendoly, E., Donohue, K., & Schultz, K. L. (2006). Behavior in operations management: Assessing recent findings and revisiting old assumptions. Journal of operations management, 24(6), 737-752.
5 Davis, A. M., Mankad, S., Corbett, C. J., & Katok, E. (2022). The Best of Both Worlds: Machine Learning and Behavioral Science in Operations Management. Available at SSRN 4258273.
6 Luong, Anh and Kumar, Nanda and Lang, Karl Reiner, Algorithmic Decision-Making: Examining the Interplay of People, Technology, and Organizational Practices through an Economic Experiment (January 31, 2020). Baruch College Zicklin School of Business Research Paper No. 2020-02-03, Available at SSRN: https://ssrn.com/abstract=3529679.
7 Boyacı, T., Canyakmaz, C., & de Véricourt, F. (2024). Human and machine: The impact of machine input on decision making under cognitive limitations. Management Science, 70(2), 1258-1275.
8 Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. Organizational Behavior and Human Decision Processes, 151, 90-103.
9 Kordzadeh, N., & Ghasemaghaei, M. (2021). Algorithmic bias: review, synthesis, and future research directions. European Journal of Information Systems, 31(3), 388–409. https://doi.org/10.1080/0960085X.2021.1927212.
10 Logg, J. M. (2019). Using algorithms to understand the biases in your organization. Harvard Business Review.
11 Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: people erroneously avoid algorithms after seeing them err. Journal of experimental psychology: General, 144(1), 114.
12 Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. American Psychologist, 34(7), 571–582. https://doi.org/ 10.1037/0003-066X.34.7.571
13 Önkal, D., Goodwin, P., Thomson, M., Gönül, S., & Pollock, A. (2009). The relative influence of advice from human experts and statistical methods on forecast adjustments. Journal of Behavioral Decision Making, 22(4), 390-409.
14 Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. Review of general psychology, 5(4), 323-370.
15 de Véricourt F, Gurkan H (2023) Is your machine better than you?You may never know. Management Sci., ePub ahead of printMay 25, https://doi.org/10.1287/mnsc.2023.4791.
16 Cao, Xinyu and Hu, Chenshan and Sun, Jiankun and Zhang, Dennis, How Forced Intervention Facilitates Long-term Algorithm Adoption (June 18, 2024). Available at SSRN: https://ssrn.com/abstract=3640862.
17 Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. Science, 185, 1124–1131.
18 Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. Review of general psychology, 2(2), 175-220.
19 Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. Science, New Series, 211(4481), 453–458.

Link to the reference list in Zotero Group Library & Contact info