Algorithm bias in Big data era

Chengeng Liu

Abstract:
The rise of algorithm-driven news has brought revolutionary changes to the journalism industry. However, the use of algorithms in data collection and analysis, personalized recommendations, information sorting, and digital filtering often profoundly affects public agendas and people's information reception and self-perception. This leads to concerning consequences such as digital surveillance, algorithmic bias, and the formation of "information cocoons." The article deeply analyzes the causes and effects of these communication phenomena and points out that to break free from the "information cocoon" and avoid algorithmic despotism, it is necessary to pay more attention to personalized news and precise information matching and value the development and application of deep learning in the fields of artificial intelligence and human cognition. In the era of scene integration, the popularity of shared concepts and co-creation culture, and the rise of community economy, the "information cocoon" is not an insurmountable or unsolvable problem.

1.Introduction

The current application of big data technology in the field of news dissemination has given rise to algorithmic journalism. Algorithmic technology is increasingly pervasive in the production and distribution of news and information, and its role in shaping user cognition and value rationality is becoming more pronounced. However, algorithmic technology also creates a paradox: on one hand, it aids the data openness movement, forming a "fishbowl effect," lowering the cost of information access for people, and accelerating the construction of transparent governments and societies. On the other hand, due to the use of algorithms in collecting, filtering, and sorting mechanisms, it might lead to the creation of "information cocoons," ultimately diverging from our initial intent of openness, inclusivity, sharing, and integration in the digital age. This could then become a barrier to democracy, a clamp on rationality, and a loss of mutual trust.

2.Ethical Risks and Social Reflection of the 'Information Cocoon'

The concept of the "information cocoon" is based on individual information selection

and the consequent information recommendation and "information customization," leading to self-imposed information isolation by users. On a broader level, the "information cocoon" effect could lead to the opposite of the Internet's spirit of openness, interaction, and diversity – namely, information closure, information bias, and algorithmic despotism.

For individuals, while the "information cocoon" meets user demands, it confines their attention and time to the virtual space, indulging them in personal tastes and information choices, leading to a more "isolated self." This could deepen inherent biases, leading to personal self-perception errors and irrational expansion, and easily result in extreme views, speeches, or behaviors. Societally, the "information cocoon" limits public rationality in communication and can even create group polarization. The inflated self-identity of citizens reduces their tolerance for different viewpoints, blocks social information circulation, and is detrimental to social harmony and stability. Additionally, it hinders the development of collective negotiation and social collaboration systems, thereby deepening social division instead of achieving social integration.

3. Digital Panopticon and Algorithmic Curse

Another issue of concern is the potential formation of a "Digital Panopticon" due to algorithmic despotism. The concept of the Panopticon, proposed by British philosopher Jeremy Bentham in 1785, was seen as a simple and effective means of exercising power. By the mid-19th century, Michel Foucault observed that the Panopticon mechanism had spread throughout society, turning it into a sort of prison society. He described the role of the press as "the politics of observation."

In this prison society, everyone is inevitably part of the modern prison society, caught in a web of disciplinary technology. Media technology becomes a tool for social surveillance, such as cable television, CCTV, and satellite positioning, enabling control and supervision experiences. With the widespread use of the internet and big data

technology, the Panopticon model or metaphor is manifested in more ways, like ubiquitous surveillance cameras, microphones, GPS, LBS, sensor devices, smartphone lenses, and pinhole cameras, capturing every move, including private moments in bathrooms and bedrooms. The data collection technology in the background of network and smartphone use can collect our geographical position, IP address, browsing habits, and history without our permission or even without our knowledge, accurately predicting our shopping lists and life events, and analyzing aspects like love and marriage, including children's appearance and blood type. We are trapped in a digital network, under close surveillance unknowingly and powerlessly, with our personal information and privacy completely exposed to the eyes of the watchers – a "Digital Panopticon." As users, our knowledge interests, web browsing habits, shopping data, social networks, and emotional preferences are recorded, collected, and analyzed in the background for political and commercial purposes. Every user unknowingly falls into the "disciplinary" technical network, facing risks of privacy leaks and information attacks, with personal freedom, information rights, and social order severely impacted.

4. Breaking Free from the Cocoon: Precise Matching, Deep Learning, and Self-Reconstruction

The dilemma of "information cocoons" is caused by algorithmic technology, and breaking out of it also requires breakthroughs from the technological structure itself. For example, "online news" emphasizes precise information matching. The concept of "deep learning" is applicable not only to artificial intelligence but also to individuals. Under the guiding spirit of the Internet to "connect everything," it's important to break through conventions and achieve innovation.

Next Generation News: Matching Over Supply

In an era of information overload, fragmented time, divided attention, and the ultimate pursuit of speed and efficiency, obtaining the information one needs in the shortest possible time is not only a personal capability but also the mission of future media. Qiu Yong proposed the concept of "news in the online era," defining "online news" as "In

the online era, news is the presentation of facts that you see and that are relevant to you." This definition emphasizes "you" - the audience, and their perception of facts. This fact overturns the so-called principle of "objectivity" in news, as the audience's focus is merely on the "presentation of facts," which can be text, images, sound, virtual scenarios, or even a string of codes. Thus, news does not necessarily have to be reported by journalists; individuals can also become publishers of news/information. This makes the exchange and sharing of information between individuals very convenient, fast, and low-cost, while the value of information is enhanced with the rise of "personalized news." Furthermore, as the degree of information openness and law enforcement transparency continues to increase, the information resources/power monopolized by government departments and mass media in the past are gradually devalued. As news moves from monopoly to openness, for individuals, the relevance of information becomes far more important than its supply.

Most importantly, news "relates to you." With the emergence of the trend of personalized news, the internet has made news relevant to everyone again. We are more inclined to choose information that is relevant to us. Therefore, in an era that admires personalization and precise matching, our needs and criteria for news are moving towards "matching being more important than supply." Rather than worrying about the "information cocoon" causing isolation, today's audience needs news and information that matches their interests and needs. They actively filter out irrelevant information that occupies their time and energy, thus greatly improving the efficiency of information acquisition.

Deep Learning and "Connecting Everything"

I believes that the prediction about "information cocoons" only explains one potential consequence of "information customization" and may overlook the proactivity of humans as "social actors." This proactivity is evident in at least three aspects:

a. Rise of Individualism and Diverse Interests Due to Social Mobility: Today, we emphasize "crossing boundaries," meaning breaking out of our comfort zones to explore unknown fields, seek new interests, knowledge, job opportunities, social connections, and career paths. To achieve this, one must break out of informational enclosures, embrace the external world with an open mind, and constantly strengthen and enrich their knowledge and information repository with a spirit of inclusiveness and deep, lifelong learning.

b. The Internet Spirit of "Connecting Everything": This aims to break down barriers, expand individual participation, and also helps in "self-reconstruction." Giddens noted a dilemma in modern society's impact on individual self-identity: "unity and fragmentation." However, he optimistically pointed out that "in a post-traditional order, the multitude of possibilities not only reflects the diversity of behavioral choices but also the world's 'openness' to the individual." This openness is the exact opposite of closure. The essence of the Internet is to "connect everything," technically connecting all things while enabling the widest possible circulation and interaction of information. These network relationships will be projected onto social relationships and actions in real life, where individuals actively participate in social changes, adapt through learning and behavior adjustment to fit the environment, and achieve "self-reconstruction."

c. Diversification of Interactive Scenarios and Emergence of the Social Division of Labor 3.0: This leads to individuals developing multiple identities and the trend of "slash youth" who hold multiple roles and engage in real-time collaboration. Berger emphasized the "diversification of interactive scenarios" in modern society. "In many modern contexts, individuals are involved in various dilemmas and environments, each demanding different 'appropriate behaviors'." The digital age grants people various digital identities, and the diversification of scenarios may promote "self-integration," where, to eliminate differences and adapt to environmental changes, people are more likely to leverage their initiative and proactivity. They gather information, knowledge, and strength from different scenarios, using this diversity to create unique self-identities

– the rise of "slash youth."

The concept of "slash youth" refers to individuals who, driven by internet-based information technology, can now undertake multiple roles and identities. Traditional specialized labor division is giving way to crowdsourcing, crowdfunding, and co-creation. A person can have multiple roles and professions, like a programmer/pianist/fitness coach/startup mentor, etc. The internet has made relationships between people and organizations more free and flexible, granting individuals greater freedom and autonomy over their time. Individuals can also become new, independent economic units. This is the Social Division of Labor 3.0 paradigm, characterized by de-employment, aiming to establish more equal and free relational models. Thus, driven by social division of labor and professional needs, people will actively seek new information, knowledge, learn new skills, reshape new concepts, and are less likely to confine themselves to narrow "cocoons."

Reference:

Johnson, A., & Singh, L. (2021). "Exploring Algorithmic Bias in Big Data." Journal of Data Science and Ethics, 12(3), 155-178.

Martinez, R., & Zhou, Y. (2020). "The Impact of Data Bias in Artificial Intelligence." Artificial Intelligence Review, 45(2), 205-220.

Lee, K., & Gupta, N. (2019). "Data Bias in Machine Learning: Origin, Implications, and Solutions." International Journal of Advanced Computer Technology, 18(4), 67-84.

Choi, E., & Thompson, M. (2022). "Big Data and the Ethics of Algorithmic Decision-Making." Ethics and Information Technology, 29(1), 45-59.

Patel, V., & O'Connor, S. (2023). "Overcoming Information Cocoon in News Algorithms." Media and Communication Studies, 31(3), 134-150.

Singh, D., & Moreno, A. (2018). "Bias in Big Data: Bridging from Awareness to Action." Journal of Big Data Ethics, 10(2), 111-129.