



SOCHUM



Topic: Addressing the racial, ethnic and gender-oriented algorithmic biases in AI technologies that are considered detrimental to the society

Committee: GA3 - Social, Humanitarian and Cultural Committee (SOCHUM)

Name: Naz Koçak

Position: President Chair

A. Introduction to the Chairboard

Dear delegates,

It is my utmost pleasure to welcome you all to the Social, Cultural and Humanitarian Committee of Kabataş Model United Nations Conference 2024. I would first like to say that it will be an honor for me to be chairing this committee. I am so excited to be a part of such a prestigious conference's chairboard. I will assure you it is going to be an amazing experience for everyone involved and I will make sure every single delegate has a productive and fun time at the conference. With that being said, I am very excited to see you all soon.

Naz Koçak

President Chair

B. Introduction to the Committee

The Committee on Social Humanitarian and Cultural Affairs (SOCHUM) of the General Assembly (GA) is one of the most significant committees in the UN system. In order to tackle global issues that fall under the categories of social humanitarianism and culture SOCHUM was established.

International norms and policies that uphold human rights combat poverty and foster intercultural understanding have been greatly influenced by it. Its extensive mandate covers a wide range of topics including disarmament, non-proliferation and peacekeeping operations. Through dialogue and negotiation, SOCHUM brings together experts from a variety of fields, civil society organizations from member states, and experts to reach a consensus on solutions to challenging global issues. With its efforts, SOCHUM significantly advances the creation of a more just equitable, and peaceful world.

C. Introduction to the Agenda Item

The use of artificial intelligence is counted as an extraordinary advancement in many areas ranging from education, health, finance and many more. The world relies on AI for many circumstances since it is perceived as unbiased and rational. However, this mysterious machinery can be abused to make unfair and biased decisions. The abuse of artificial intelligence offers significant dangers in a variety of fields, ranging from growing societal disparities to compromising world security. An immediate problem arises within the persistence of the biases inherent in AI systems. When educated on biased datasets or constructed without sufficient observation, AI systems can unintentionally reinforce existing societal biases. This can result in discriminatory consequences in critical areas including hiring practices and criminal justice proceedings, further marginalizing already vulnerable groups.

The widespread use of AI in a variety of fields also pose a question of privacy. Governments and private organizations rely heavily on AI for uses like widespread surveillance, facial recognition, and predictive analytics, thus furthering concerns regarding individual autonomy and privacy. The unrestricted development of surveillance capabilities, paired with inadequate measures to preserve personal information, highlights the necessity of universal organization and observance to defend privacy in the digital era.

D. Key Terms

a. Artificial Intelligence

Artificial intelligence, commonly known as AI, is the ability of a computer or a machine to demonstrate human-like tasks such as problem-solving, identifying and remembering past actions to determine the next (machine learning). It is done by taking a piece of data, processing it and learning from past actions, without any form of human interference.

b. Algorithm

An algorithm is a basic set of rules designed to be followed by a machine, such as a computer, to solve a problem presented to it. It is the key component of the machine since it is the driving force behind it.

c. Machine Learning

Machine learning is an area of artificial intelligence that is about analyzing data and learning from it to make similar decisions in the future. With machine learning, the computer doesn't need explicit directions on doing the task it's given since it already generalized it before and knows what to do.

d. Deep Learning

Deep learning is a subset of machine learning that includes layered data to make complex and more human-like decisions. While machine learning technologies need manual algorithms to initiate, deep learning robots can determine these algorithms by themselves.

E. General Overview

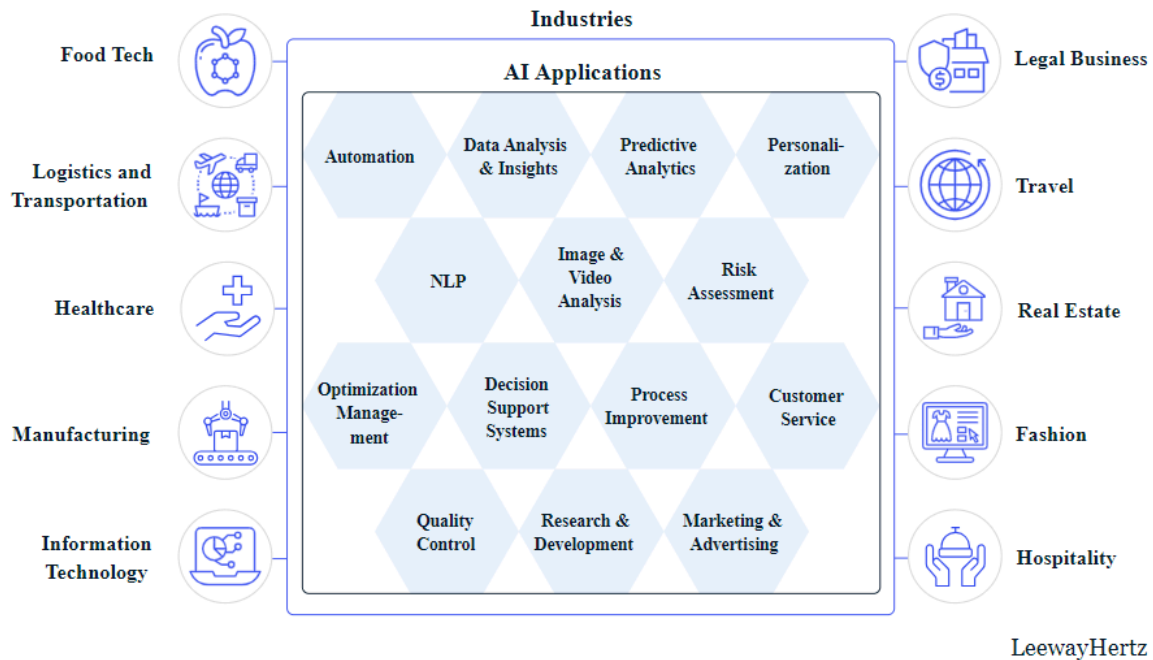
Artificial intelligence systems try to imitate tasks done by humans, such as learning and understanding. They perform their functions by processing data and gathering an output with the information they received from external resources. Machine learning technologies also have the ability to predict future outcomes. Therefore, already existing information is crucial to this machinery.

AI works with an algorithm which fundamentally shapes the decision making process. It is through these algorithms that make the AI recognize patterns and past actions. There are many types of learning algorithms the robot can use, one of the most important one being machine learning. The robot's ability to recognize makes it 'smarter' for most instances. For example, if a picture of a cat is introduced to the machine, with the algorithm it possesses, it will remember what cats look like and won't need any future guidance. But this can also be a problem if the algorithm doesn't have enough data to distinctly separate the features of a cat from other things. With the lack of previous knowledge, it will mislabel and give inaccurate results. Another subcategory of artificial intelligence that increases the accuracy of the results of the machinery is deep learning. Deep learning has gained popularity over the past few years, since its technology can process larger amounts of data and recognize more complex patterns in text, speech or images.

These types of algorithms enable a framework which demonstrates that they have made an error, and this error is transcribed to the existing system so that it cannot recur. Even though this process guarantees more accurate results for the upcoming conducts, the artificial intelligence still requires human interference and fails to recognize the error that was made without it being pointed out.

a. Fields in Which Artificial Intelligence (AI) Technology is Used

Artificial intelligence has proved that it has the potential to better lives in many ways. Its effect goes beyond basic needs and appears to enhance the human experience overall. With this quality, artificial intelligence systems are used in many fields.



These areas of work rely now on artificial intelligence systems in some shape or form, whether if it is to analyze a DNA pattern in a research center or to estimate the income of a financial department. The fields that are affected from the biased and stereotyping artificial intelligence the most are crime prevention and employee recruitment fields. Especially facial recognition systems play a role in the biased systems.

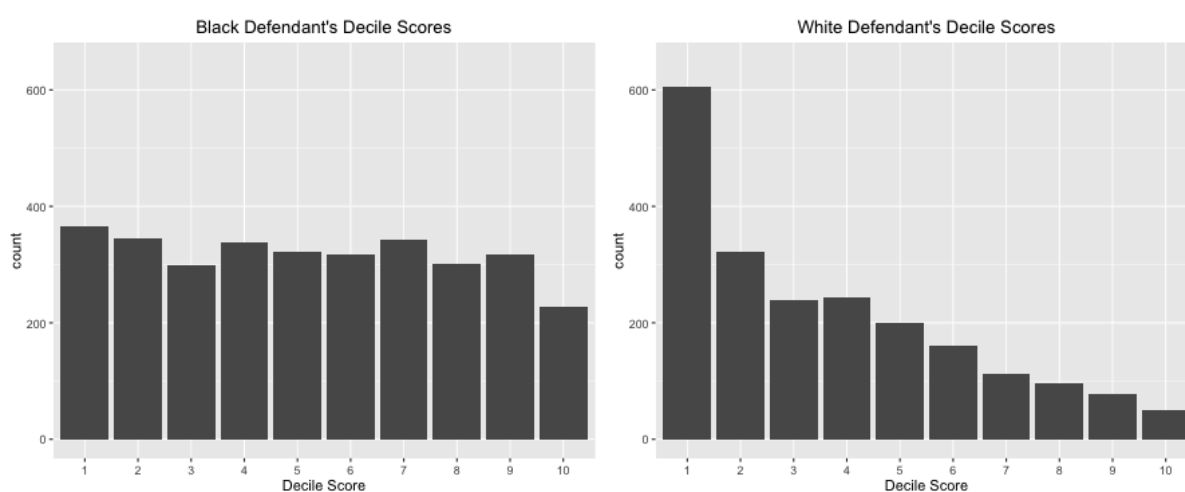
b. Bias in Artificial Intelligence (AI) Algorithms

It is rare for engineers to purposefully inject biases into algorithms, and algorithms are not inherently discriminatory. However, automated recruitment is still open to biases. The underlying technologies of artificial intelligence and machine learning are intimately related to this problem. The three main steps in machine learning that contribute to algorithmic biases are feature selection, target formulation by the engineers, and dataset generation. These steps can be broken down into smaller phases. When editors handle unstructured data and engineers construct algorithmic rules, biases may be introduced since the dataset lacks diverse representation from different organizations.

One of the reasons for biases in an algorithm is lack of data. If a machinery lacks the crucial pieces of information it needs, its system will fail to make the necessary rational and fair decision. This problem occurs when datasets are biased toward more mainstream and easily accessible groups. As a result, the distribution of gender and racial dimensions is unbalanced. In the event that a specific race or gender is not sufficiently represented in the collected data, the resulting system will unavoidably ignore them or treat them unfairly in its operation.

c. Racial Bias in Artificial Intelligence (AI)

One of the most dangerous fields artificial intelligence causes a risk to is crime prevention. The use of data in police work is not a new trend, however, with the introduction of artificial intelligence to this area, it has taken a new turn. The Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) is a software developed to assess if a defendant is likely to repeat a criminal behavior. The output given by the system helps the judges come to a verdict. COMPAS has been used by the U.S. states of New York, Wisconsin, California, Florida, and other jurisdictions in the country. Skin tone or race are not accepted as data by the COMPAS system. However, a 2016 study by ProPublica revealed that COMPAS is "biased against blacks."



ProPublica summarizes: "COMPAS correctly predicts recidivism 61 percent of the time. But blacks are almost twice as likely as whites to be labeled a higher risk but not actually reoffend. It makes the opposite mistake among whites. They are much more likely than blacks to be labeled lower risk but go on to commit other crimes. Black defendants were also twice as likely as white defendants to be misclassified as being a higher risk of violent recidivism. And white violent recidivists were percent more likely to have been misclassified as a low risk of violent recidivism, compared with black violent recidivists."

d. Gender Bias in Artificial Intelligence (AI)

Another sector in which using artificial intelligence can be dangerous is the hiring field. It is more common for artificial intelligence to pose gender discriminatory behavior in this area of work. In 2014, a machine learning based hiring algorithm was created by Amazon for analyzing job applications. The system did not classify candidates on a gender-neutral basis. The bias resulted from training the artificial intelligence system on primarily male personnel CVs. As a result, the recruitment algorithm interpreted this biased model as an indicator of success, leading to prejudice

against female applicants. The program even downgraded applications that included terms like "female". These discoveries pushed Amazon to withdraw the service and create a new, unbiased algorithm. This prejudice was unintentional, nevertheless showing the weaknesses inherent in algorithmic bias, which reinforces existing gender inequities and social biases.

e. Skin Tone Bias in Artificial Intelligence (AI)

In 2017, Google's photo searching algorithm mistakenly labeled a photo of two black men as gorillas since the algorithm didn't have enough data to distinguish between different skin tones. Quickly after this accident, Google apologized and took immediate measures to prevent a mistake like this happening ever again. Three years later, the company discontinued the picture identification service, stating that they have to address technical issues in order to continue the service. Similarly, also in 2017, a contactless soap dispenser's algorithm failed to correctly distinguish skin tones, causing the dispenser to only respond to white hands and not detect black and brown ones. These situations demonstrate algorithmic biases.

F. Past Actions

United Nations Educational, Scientific and Cultural Organization (UNESCO) became the first United Nations agency to work on the gendering of artificial intelligence with the release of a publication in March 2019 entitled 'I'd Blush if I Could'. This publication featured ideas on actions to overcome global gender gaps in digital skills, with a special examination of the impact of gender prejudice coded into some of the most common artificial intelligence applications. The paper links this prejudice to the gender imbalance of technical teams developing new and advanced technologies, which frequently results in the reproduction of prejudices by robots. Since its introduction, "I'd Blush If I Could" has helped start a global discourse on gendering.

On November 23, 2021, UNESCO's Member States accepted the first global legislative document on the ethics of artificial intelligence. The document addresses concerns raised in the Human Rights Council Resolution about AI. UNESCO's Recommendation on the Ethics of Artificial Intelligence is an important agreement among Member States on the values, principles, and rules that should guide the development of this transformative technology.

In September 2021, the United Nations Institute for Training and Research (UNITAR) launched an online training course for youth on artificial intelligence and human rights. The course covers the impact of artificial intelligence on freedom of expression, privacy, equality, and challenges such as filter bubbles, echo chambers, automated content moderation, and data-driven biases. This course allows young people to gain knowledge on the topic, which arguably may be the most effective way of preventing future biases in algorithmic systems.

G. Guiding Questions

- a.** What are the particular human rights implications of artificial intelligence technologies like facial recognition, predictive policing, and algorithmic decision-making, and how do they affect disadvantaged or vulnerable communities?

- b.** How do international organizations, such as the United Nations and regional agencies, help to shape norms and standards for responsible artificial intelligence governance and to protect human rights?

- c.** What role do nonprofit organizations, educational institutions, and the private sector play in promoting artificial intelligence development and deployment practices that prioritize transparency, accountability and fairness?

H. Bibliography

Turing, A. M. "I.—COMPUTING MACHINERY AND INTELLIGENCE." *Mind*, vol. LIX, no. 236, Oct. 1950, pp. 433–60. <https://doi.org/10.1093/mind/lix.236.433>

Discrimination, Artificial Intelligence, and Algorithmic,
rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73.

Team, Investopedia. "What Is Artificial Intelligence (AI)?" *Investopedia*, 9 Apr. 2024,
www.investopedia.com/terms/a/artificial-intelligence-ai.asp.

What Is Artificial Intelligence (AI)? | IBM. www.ibm.com/topics/artificial-intelligence.

Chen, Zhisheng. "Ethics and Discrimination in Artificial Intelligence-enabled Recruitment Practices." *Humanities & Social Sciences Communications*, vol. 10, no. 1, Sept. 2023,
<https://doi.org/10.1057/s41599-023-02079-x>.

What Is Deep Learning? | IBM. www.ibm.com/topics/deep-learning.

UNESCO Recommendation on the Ethics of Artificial ...,
www.ohchr.org/sites/default/files/2022-03/UNESCO.pdf.