

Natalia Jara ¹

'Discrimination is not just human'

Introduction

The pandemic will leave us with a new perception of the benefits of Artificial Intelligence (AI) technologies for decision making. All over the world people are acknowledging that this technology has been used on a daily basis without their knowledge. But they are also recognising that its implementation has become a necessity to the public interest.

Accordingly, governments and businesses have deployed different types of AI to fight the pandemic. They have improved diagnoses or developed treatments. However, what if someone is banned from public transport because an AI predictive model marks them as a potential source of infection? What if someone is refused health insurance or a mortgage because of a Covid-19 infection? Are results obtained by an AI enough to justify restricting individual freedom?

From a legal perspective, concepts like "discrimination" or "objective" are challenged when AI technology gets involved in decision making. AI-assisted decisions are based on algorithms or maths. This, arguably, gives them some level of objectivity. But their outcomes could nonetheless be discriminatory.

What is artificial Intelligence decision-making?

Artificial Intelligence is defined in the Oxford English Dictionary as "the theory development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages".

AI uses algorithms to classify, analyse, create profiles or make predictions to perform tasks that normally require human intelligence. This includes selecting employees or students, advertising, evaluating insurance risks, image search analysis and preventing the spread of a pandemic.

Algorithms, in turn, are described by Frederik Zuiderveen Borgesius, professor of law at the Radboud University Institute for Computing and Information Sciences in the Netherlands, as "a documented series of steps which leads to the transformation of some data". He calls them "an abstract, formalized description of a computational procedure". They are instructions that guide the decision-making processes, step by step, to obtain a certain result or solution.

AI and algorithms are thus closely related concepts. An algorithm is the internally coded instructions that allow an AI to work and fulfill its purpose.

With the ease and accuracy that algorithms and AI can organise large amounts of data, this technology could improve any government, organization or individual's decision-making process. In fact, any social, economic or political decisions that involve a combination of human and AI systems could be

considered partially or wholly automatic.

An algorithm's true nature

Nevertheless, are these algorithms completely free of prejudice and discrimination? Can we really rely on algorithms or AI to make objective decisions? Are they not just machines that process data with logic and math?

The truth is that many cases around the world have proved that algorithms and AI are not neutral, and free of neither prejudice nor discrimination. There have been algorithms which recognise certain people as “gorillas”. Others display better paid jobs to men than to women or even discard women’s resumes during job applications. In yet another case, people living in dangerous neighborhoods have been prevented from paying for express delivery services. These are just a few examples of people being treated differently simply because of their gender or phenotype.

If society wants to rely on this technology to improve the decision-making process, it is necessary to comprehend how these prejudices are transferred to or adopted by algorithms and AI.

Algorithmic and AI Discrimination

In the words of Ramón López de Mántaras, director of the Instituto de Investigación en Inteligencia Artificial-CSIC, in Barcelona: “If the data you learn from is biased (intentionally or not), the algorithm will decide biased”.

Algorithms are simply instructions for a computer procedure. The data that it receives as input could replicate or even amplify a human prejudice or bias. For this reason, AI could perpetuate discriminatory data analysis processes and further predictions that could worsen various existing stereotypes in society.

If so, can the risk of discrimination through fully or partly automatic making-decision processes be reduced or mitigated if the data collecting process is done carefully and correctly?

Experts, like MIT Technology Review writer Karen Hao, say that this is not going to be enough. The algorithm itself could lead to a prejudiced decision if programmed to analyse the data in an “incorrect” or “biased” way. “Historical patterns of discrimination”, Hao argues, are discussed by AI designers since this technology could learn how to think just like a human being. This includes imitating our prejudices or stereotypes.

On the other hand, we should consider that no AI system currently has its own intentionality, yet. So, any biased decision taken in a partial or fully automatic decision-making process that results in discrimination is down to bias in the imputed data or the processing programming. Either way, the intention or bad faith must be related to the programmer or AI system proprietor.

With this in mind, when evaluating an AI system or algorithm, its decisions must be analysed from a legal perspective to identify cases of discrimination.

How do we protect ourselves from AI-driven discrimination?

As we all know, discrimination is prohibited by many treaties and constitutions around the world. But

will this be enough to protect and secure equal treatment when decisions are made by an AI system?

Article 7 of the UN's Universal Declaration of Human Rights reads:

“All are equal before the law and are entitled without any discrimination to equal protection of the law. All are entitled to equal protection against any discrimination in violation of this Declaration and against any incitement to such discrimination”.

Accordingly, all people are entitled to not suffer any violation of their human rights caused by an AI driven discriminatory system or algorithm. However, this will depend on the efficiency of each country's laws or regulations. There are numerous approaches to accomplishing a non-discriminatory automatic algorithm or AI- augmented decision-making process.

On one hand, in the European Union, Article 14 of the European Convention of Human Rights (EHRC) prohibits discrimination by saying:

“The enjoyment of the rights and freedoms set forth in this Convention shall be secured without discrimination on any ground such as sex, race, color, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status”.

Furthermore, the EHRC distinguishes between direct and indirect discrimination. Direct discrimination targets a person based on a particular characteristic, such as a ban against people of a certain religion. If discrimination occurs as a result of apparently neutral practice – like a regulation against all head coverings that, by default, prevents observant religious people from applying for a job – it might be considered to be indirect.

Both types of discrimination could be caused by an algorithm or an AI-driven system, especially the indirect type. Algorithms and AI systems can unintentionally disadvantage a certain group of people of a particular phenotype, culture or gender, through seemingly neutral criteria, practices or selections. For a recent case illustrating this, see the European Court of Human Rights case ECtHR, Biao v. Denmark (Grand Chamber), No. 38590/10, 24 May 2016, paragraphs 91 and 92.

Consequently, current European regulations require that general policies, measures or practices based on algorithms or AI system are analysed and perhaps approved prior to use, since it is unclear whether they might breach discrimination prohibitions.

The American way

On the other hand, considering that the focus in European regulation is not on intentions but on whether a discriminatory decision is made without an objective or reasonable justification, the problem of lack of regulation of “black boxes” will be important. “Black boxes” refer to the internal programming of an AI system. These programs are often kept secret by IT companies, but would need to be transparent to allow a judge to analyse how potentially discriminatory decisions were made.

Situations similar to those in the EU will occur on Latin American countries like Argentina, Bolivia, Chile, Ecuador, Mexico, Peru and Uruguay. In most of their constitutions and regulations require equal rights and prohibit discrimination. But discovering whether an unequal treatment was “arbitrary” or “unreasonable” is necessary to find whether a decision or act is discriminatory. Therefore, in this region, the question of how to make an AI systems' black box process transparent to justify acts or

decisions assisted by an algorithm will be significant. So will the issue of how to prevent biased data from being processed.

Finally, another interesting approach is found in the United States, where “discrimination” in Barron’s Law Dictionary is defined as:

“the unequal treatment of parties who are similarly situated. Federal law prohibits discrimination basis of race, sex, nationality, religion, and age in matters of employment, housing, education, voting rights and access to public facilities. Furthermore, states or any governmental bodies may not engage in any actions which result in discrimination on the grounds of race, sex, nationality, religion or age”.

According to this definition, importance would be placed in the possible engagement in any action that may result in discrimination. In this sense, if any state or governmental body engages or uses a discriminatory AI-driven technology, it would be illegal even though it does not result in an unequal treatment. However, if any private organization uses AI-driven technology, the illegal action will only occur, if the result is discrimination.

Conclusion

Partially or fully automatic making-decision processes assisted by algorithms or AI systems could cause, and even amplify, bias and discrimination. Regulations all over the world will be challenged by the increasing use of this type of technology in decision making.

Topics like biased data processing, black box transparency, intentionally discriminatory programs or policies and measures to supervise algorithms or AI technology must be discussed and improved. This is particularly true since urgent situations, like a worldwide pandemic, drive the use of increasingly intelligent, accurate and faster technology.

Nonetheless, for the moment, government bodies and businesses which use this technology to ease their decision-making processes will have to design measures to provide transparency of how their algorithms run or justify outcomes in a reasonable way. As we have seen, our human rights are at stake.

¹ Natalia Jara is Associate Attorney.