

The risk of discrimination by algorithm

November 13 2019



Not always fair: When humans are evaluated by algorithms, care must be taken.
Credit: Patrick Langer, KIT

Not only companies but state institutions increasingly rely on automated decisions by algorithm-based systems. Their efficiency saves time and money, but also entails many risks of individuals or population groups being discriminated against. This is the result of a study made by the Institute for Technology Assessment and Systems Analysis (ITAS) at

Karlsruhe Institute of Technology (KIT) on behalf of the Federal Anti-Discrimination Agency.

When granting a loan, selecting new staff members, or making legal decisions—in an increasing number of sectors, algorithms are applied to prepare human decisions or to make these decisions for humans.

"Unfortunately, it often is a mistake to think that this inevitably leads to more objective and fairer decisions," says Carsten Orwat of the Institute for Technology Assessment and Systems Analysis (ITAS) of KIT.

"Situations become particularly critical when algorithms work with biased data and rely on criteria that ought to be protected," the author says. These criteria include, in particular, age, gender, ethnic origin, religion, sexual orientation, and handicaps.

On behalf of the Federal Anti-Discrimination Agency, Carsten Orwat studied in detail the causes of discrimination, its impact on society, and future options to reduce discrimination risks. His study entitled "Diskriminierungsrisiken durch Verwendung von Algorithmen" (discrimination risks by using algorithms) lists 47 examples to illustrate how algorithms can discriminate against people in various ways and how this can be detected and proved.

Real Estates, Loans, Judicial Matters, and More: Various Examples of Discrimination Risks

As examples, Orwat describes situations on the [real estate](#) and loan markets or in the court system. In the USA, for instance, several cases have been documented, in which algorithms within social media permitted targeted advertisements to be invisible to persons protected by the "Fair Housing Act," such as migrants, people with handicaps, or with non-white skin color, the author says. In Finland, a bank was sentenced to pay a fine because its [algorithm](#) for the automatic granting of online

loans showed bias towards men over women and Finnish over native Swedish speakers. This unequal treatment is forbidden by Finnish anti-discrimination law. When deciding on early releases from prison, US judges use a much disputed system that calculates risk scores. Journalists and human rights associations criticize the fact that this system systematically overestimates black people's risk of re-offending.

"Machine learning systems often have problems when they are trained with data reflecting unequal treatments or stereotypes," Carsten Orwat explains. "In this case, the algorithms generated will also do so. When processing data containing evaluations of people by other people, unequal treatments and discriminations may even spread or increase." This happened in the USA in a system for food and health controls that was based on discriminating ratings of restaurants.

Recommendations of Countermeasures

However, society must no longer accept these unequal treatments. The study lists several options to counteract discrimination by algorithms. "Preventative measures appear to be most reasonable," Carsten Orwat says. Companies may ask anti-discrimination agencies to instruct their staff and IT experts and increase their awareness. Then, these persons will use datasets that do not reflect any discriminating practices or unequal treatments.

According to Orwat, the goal is to make future algorithms "discrimination-free by design." This means that programs have to be checked during their initial development.

In the end, it is all about the protection of the society's values, such as equality or free development of the personality. To guarantee this in spite of the very rapid developments of "big data" and AI, it is necessary to improve anti-discrimination and data protection legislation at some

points, Orwat points out.

Provided by Karlsruhe Institute of Technology

Citation: The risk of discrimination by algorithm (2019, November 13) retrieved 26 April 2024 from <https://techxplore.com/news/2019-11-discrimination-algorithm.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.