Principles: Safety and security, Fairness and non-discrimination, Human oversight and determination, Transparency and explainability, Responsibility and accountability, Multi-stakeholder and adaptive governance and collaboration
Values: Human Rights
Stakeholders: Civil society, Public sector, Technical community

Biased Recidivism Assessment from Algorithms

The most recent developments in AI indicate that algorithmic predictions are more accurate than human authorities when comes to calculating the risk of recidivism for offenders, with algorithms clearly outperforming their human counterparts. A software called ‘‘Correctional Offender Management Identification Tools’ (COMPIT) was developed by a private company in the USA to assist human judges during the sentencing stage.

The aim of COMPIT is to deliver more accurate, unbiased results where their human counterparts fail to do so. While Northpointe Inc. assured their clients COMPIT predictions were in the “good range of predictive accuracy”.

There were findings in 2016 that African-American offenders assessed by the assessment tool were two times more likely to be given a recidivism high-rate risk than the rest of the population. The risk assessment software is comprised of 137 factors such as age, gender, previous criminal history, and omits race. COMPIT is used to calculate the risk score of recidivism for offenders on a scale of 1 to 10, 1 being less likely, 10 being most likely).

One particular defendant of African ethnicity was being assessed by the AI tool and subsequently given a much higher rate than his white counterpart who had committed the same crime; they had both stolen approximately $80 worth of supplies from an Office Depot store.

Algorithmic bias is a common concern when evaluating the ethical implications of AI. Bias can be easily inherited by machines from their creators and can also be a result of unintentionally including proxy information such as postal codes or addresses in training data sets. Discriminatory bias in AI being used in criminal justice needs to be addressed at the root of its developmental stages. The private entities developing algorithms being used in a public driven sector such as the judicial sector should also disclose the datasets being implemented.

- Developers of Risk Assessment Instruments (RAI) should acknowledge and adapt their training data to make it as inclusive as possible by using an intersectional approach to their hiring process.

- Developers of RAIs should also make their development processes of algorithms that are for public use as transparent as possible so that they can be held accountable for the data they acquire and use.

- Judiciary members such as judges or parole board members should ensure that there is always a human component involved with any kind of decision-making process and consider other factors when making their decisions.

- Policy makers need to be aware of how these algorithms currently being used in the judicial sector and are capable of facilitating and reinforcing the current inequalities embedded in society.

Responsibility & accountability Fairness and non-discrimination, Respect and protection of human dignity, safety, and security.

Know more about this case:

- “Can AI reduce recidivism.”, GCN, https://gcn.com/articles/2020/08/13/ai-reduce-recidivism.aspx

- “Practitioner’s Guide to COMPAS”, Northpointe, http://www.northpointeinc.com/files/technical_documents/FieldGuide2_081412.pdf

- “The Impact of Artificial Intelligence on the Right to a Fair Trial: Towards a Robot Judge?”, Asian Journal of Law and Economics, https://www.researchgate.net/publication/343258658_The_Impact_of_Artificial_Intelligence_on_the_Right_to_a_Fair_Trial_Towards_a_Robot_Judge

- "Machine Bias", Pro Publica, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing