

Com Grandes Dados, Vem Grande Responsabilidades

Big Data é um termo que tem sido uma tendência ao redor do mundo por prover, através de uma grande quantidade de dados (daí o termo), meios para entender e avançar em conteúdos de diversas áreas. Algoritmos e modelos matemáticos têm sido utilizados para através desses dados aprimorar processos como: contratações para empregos, definição da quantidade de crédito dado a uma pessoa, admissão em universidades. Porém, estes algoritmos e modelos matemáticos podem se tornar “Armas de Destruição Matemática”, ou WMD (*Weapons of Math Destruction*) como Cathy O'Neil cita em seu livro “Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy”. WMD são soluções matemáticas e/ou computacionais que visam otimizar processos baseando-se em dados previamente adquiridos, mas que acabam criando desigualdade e até discriminação através de fatores como raça, sexo, religião, orientação sexual, deficiência, etc.

Pode-se citar como exemplos de casos reais, pessoas serem rejeitadas a ganhar crédito no banco simplesmente por morar em bairros pobres, mulheres serem rejeitadas de um emprego por haver majoritariamente homens na empresa, ou até mesmo negros que não são contratados por não haver uma grande representação de negros nos dados. Este tipo de comportamento pode levar a um ciclo social discriminatório. Se mulheres não são contratadas para um determinado cargo, logo haverá uma redução de mulheres trabalhando nesta determinada área no mercado, fazendo com que os dados continuem sendo tendenciosos a escolha de homens, gerando mais desigualdade. Esses tipos de acontecimentos mostram que, para se fazer uso dessas técnicas, elas devem ser pensadas de forma que haja garantia de que não haverá nenhum tipo de enviesamento.

Há diversos desafios ao tratar dados de forma automática, para que elas não se tornem armas de destruição social, gerando injustiça. Dois dos maiores desafios são: *inputs*, que tem relação com os dados em si, e a construção das técnicas de automatização. Um bom algoritmo de uma grande massa de dados necessita processos sofisticados, e esses processos precisam de dados bem definidos. Erros na entrada de dados como: escolhas pobres de parâmetros nos dados, dados incompletos ou errados, dados que não representam a população que será avaliada e dados que são baseados exemplos antigos, podem fazer com que a desigualdade e preconceito só aumentem. Além de uma possível falha por causa dos dados, a falta de transparência nesses algoritmos, no sentido de receber uma entrada e devolver uma saída sem conhecer o que está por trás da decisão, torna ainda mais difícil o reconhecimento de possíveis falhas, como: correlacionar variáveis que não deveriam ser correlacionadas, admissão de um certo tipo de comportamento que não deveria existir, entre outras.

O uso de *big data* tem a chance de revolucionar o jeito que convivemos socialmente e aumentar a qualidade de vida, porém, para que isso seja possível, é imprescindível que haja estudos, discussões, acordos e novas propostas, para que estas técnicas que usam os dados para otimizar processos sociais, não afetem determinados grupos. É então necessário, que setores governamentais e privados atuem em conjunto para que os valores justos, valores contra discriminação, preconceito e desigualdade, sejam depreciados por qualquer tipo de sistema que utilize dados para gerar oportunidades entre as pessoas. Como O'Neil cita no último capítulo de seu livro: "A tecnologia já existe. É apenas a vontade que nos falta.", mostra que não está sendo feito tudo que pode ser feito, ou seja, que com um esforço maior das entidades responsáveis, a *big data* pode ser usada para trazer mais igualdade e justiça para a execução de diversos processos.