

## **A Era da Fé Cega em Big Data Deve Acabar | Cathy O'Neil**

Os algoritmos estão em todo o lado. Classificam e separam vencedores de perdedores. Os vencedores conseguem o emprego ou uma boa oferta de cartão de crédito. Os perdedores nem sequer conseguem uma entrevista ou pagam mais pelo seguro. Estamos a ser avaliados por fórmulas secretas que não entendemos e que, muitas vezes, não têm sistemas de recurso. Isso levanta a questão: e se os algoritmos estiverem errados?

Para construir um algoritmo, é necessário duas coisas: dados, ou seja, o que aconteceu no passado, e uma definição de sucesso, a coisa que se procura e frequentemente se deseja. Treina-se um algoritmo através da observação e descoberta. O algoritmo descobre o que está associado ao sucesso. Que situação leva ao sucesso?

Na verdade, todos utilizam algoritmos, apenas não os formalizam em código escrito. Deixe-me dar um exemplo: eu uso um algoritmo todos os dias para preparar uma refeição para a minha família. Os dados que utilizo são os ingredientes na minha cozinha, o tempo disponível e a minha motivação, e escolho os dados que considero relevantes. Não considero aqueles pequenos pacotes de massa instantânea como comida. (Risos) A minha definição de sucesso é: uma refeição é bem-sucedida se os meus filhos comerem legumes. A definição seria muito diferente se o meu filho mais novo estivesse no comando. Para ele, o sucesso seria poder comer muita Nutella. Mas sou eu quem escolhe o sucesso. Eu estou no comando. A minha opinião importa.

Esta é a primeira regra dos algoritmos: os algoritmos são opiniões codificadas. Isto é muito diferente do que a maioria das pessoas pensa sobre os algoritmos. Pensam que os algoritmos são objetivos, verdadeiros e científicos. Isso é uma estratégia de marketing. Também é uma forma de marketing intimidar-nos com algoritmos, levando-nos a confiar e a temer os algoritmos porque confiamos e tememos a matemática.

Muita coisa pode correr mal quando depositamos uma fé cega em big data. Esta é a Kiri Soares, diretora de uma escola secundária no Brooklyn. Em 2011, ela contou-me que os seus professores estavam a ser avaliados por um algoritmo complexo e secreto chamado “modelo de valor acrescentado”. Eu disse-lhe: “Bem, descobre qual é a fórmula e mostra-me. Eu explico-te.” Ela respondeu: “Tentei obter a fórmula, mas o meu contacto no Departamento de Educação disse-me que era matemática e que eu não a iria compreender.”

A situação piorou. O *New York Post* fez um pedido ao abrigo da Lei de Liberdade de Informação, obteve os nomes de todos os professores e as suas pontuações e publicou-as como um ato de humilhação dos professores. Quando tentei obter as fórmulas e o código-fonte pelos mesmos meios, disseram-me que não podia. Fui recusada. Mais tarde, descobri que ninguém na cidade de Nova Iorque tinha acesso a essa fórmula. Ninguém a entendia.

Então, alguém muito inteligente, Gary Rubinstein, decidiu intervir. Ele encontrou 665 professores dos dados do *New York Post* que tinham duas pontuações, algo que acontecia se ensinassem matemática do 7.º e do 8.º anos. Ele decidiu representá-las num gráfico. Cada ponto representa um professor. (Risos) O que é isto? (Risos) Isto nunca deveria ter sido usado para avaliações individuais. É quase um gerador de números aleatórios. (Aplausos)

Mas foi usado. Esta é a Sarah Wysocki. Ela foi despedida, juntamente com outros 205 professores, do distrito escolar de Washington, DC, apesar de ter ótimas recomendações do seu diretor e dos pais dos seus alunos. Sei o que muitos de vocês estão a pensar, especialmente os cientistas de dados e especialistas em IA. Pensam: “Eu nunca criaria um algoritmo tão inconsistente.” Mas os algoritmos podem falhar e até ter efeitos profundamente destrutivos, mesmo com boas intenções.

Enquanto um avião mal projetado cai e todos o veem, um algoritmo mal projetado pode continuar a causar danos silenciosamente durante muito tempo.

Este é o Roger Ailes. (Risos) Fundou a Fox News em 1996. Mais de 20 mulheres denunciaram-no por assédio sexual. Disseram que não tinham permissão para prosperar na Fox News. Ele foi expulso no ano passado, mas os problemas persistiram.

Isso levanta a questão: o que deveria a Fox News fazer para mudar de página? E se substituíssem o processo de contratação por um algoritmo de aprendizagem automática? Parece uma boa ideia, certo? Pensem nisso. Quais seriam os dados? Uma escolha razoável seriam os últimos 21 anos de candidaturas à Fox News. E a definição de sucesso? Talvez alguém que tenha ficado lá durante quatro anos e tenha sido promovido pelo menos uma vez. Parece razoável.

O algoritmo seria então treinado para identificar padrões de sucesso com base nesses dados. Mas se aplicássemos isso a um grupo atual de candidatos, o algoritmo filtraria as mulheres, porque não se parecem com as pessoas que tiveram sucesso no passado. Os algoritmos não tornam as coisas justas se os aplicarmos cegamente. Eles apenas repetem as nossas práticas passadas e automatizam o status quo.

Isto seria ótimo num mundo perfeito, mas não vivemos num mundo perfeito. Muitos cientistas de dados são instruídos a seguir os dados e a focar-se na precisão. O problema é que todos temos preconceitos, e esses preconceitos podem ser codificados nos algoritmos.

Façamos um exercício mental: imagine uma sociedade totalmente segregada, onde a polícia é enviada apenas para os bairros minoritários para procurar crimes. Os dados sobre prisões seriam enviesados. Se pedíssemos a cientistas de dados para prever onde ocorreria o próximo crime, a resposta seria “num bairro minoritário”.

A realidade não é tão drástica, mas temos segregações severas em muitas cidades, policiamento tendencioso e dados judiciais enviesados. Algoritmos de previsão de reincidência são um exemplo. A organização *ProPublica* analisou um desses algoritmos usado na Florida. Bernard, um homem negro, foi classificado

como de alto risco. Dylan, um homem branco, foi classificado como de baixo risco, apesar de ter um crime grave no histórico.

O que está a acontecer? Lavagem de dados. Escondem verdades desagradáveis dentro de algoritmos de caixa preta e chamam-lhes objetivos e meritocráticos. Eu chamo a estes algoritmos “*armas de destruição matemática*”. (Risos) (Aplausos)

Estes algoritmos estão por todo o lado, criados por empresas privadas para fins privados. Mesmo quando são vendidos a instituições governamentais, invocam o “segredo comercial” para evitar a transparência.

O mercado não resolverá isto. Há muito dinheiro a ser feito com a injustiça. Todos somos tendenciosos de formas que nem reconhecemos. Isso reflete-se nos dados que escolhemos e na definição de sucesso que aplicamos.

Precisamos de auditar os algoritmos para garantir justiça. Os algoritmos podem ser questionados e corrigidos. Chamo a isto uma *auditoria algorítmica*. Devemos verificar a integridade dos dados, questionar a definição de sucesso e considerar os erros e os efeitos de longo prazo dos algoritmos.

Cientistas de dados não devem ser árbitros da verdade. Devem ser tradutores de discussões éticas. E para o resto de vocês: isto não é um teste de matemática, é uma luta política. Precisamos de exigir responsabilidade dos nossos senhores algorítmicos. (Aplausos)

A era da fé cega em big data deve acabar. Muito obrigado. (Aplausos)