

Discriminação algorítmica: e o problema do ruído?

Porque a mera remoção do ruído não evita julgamentos discriminatórios e incorretos

Parte XIII

Ana Frazão

Advogada. Professora de Direito Civil e Comercial da UnB. Ex-Conselheira do CADE.

No primeiro artigo da presente série, mencionei a opinião de Daniel Kahneman – antecipando que a considerava excessivamente otimista - em relação à defesa de fórmulas e julgamentos algorítmicos. Ao longo dos artigos, procurei demonstrar, com base na opinião de vários autores, algumas das razões pelas quais os julgamentos algorítmicos podem conter diversos problemas e apresentar graves riscos, incluindo os de discriminação.

Neste artigo, voltarei a Kahneman, mais especificamente à sua nova obra em coautoria com Sibony e Sunstein: *Noise. A Flaw in Human Judgment*¹. Afinal, neste livro, tem-se mais uma contribuição que realça a importância dos julgamentos algorítmicos, mas, pelo menos ao meu ver, sem levar em consideração todos os seus riscos e ainda partindo de algumas premissas que são bem discutíveis.

Para situar o leitor nesse debate, é importante definir inicialmente o que seriam os ruídos. Para os autores, são erros não sistemáticos, devidos a fatores imprevisíveis, mas que podem impactar alarmante nos julgamentos humanos. Sob essa perspectiva, os ruídos diferenciam-se dos vieses, porque

¹ KAHNEMAN, Daniel; SIBONY, Olivier; SUNSTEIN, Cass. *Noise. A Flaw in Human Judgment*. Little Brown Spark, 2021.

estes últimos são sistemáticos e, conseqüentemente, previsíveis. Já ruídos apenas podem ser identificados estatisticamente.

Logo, ruídos e vieses são diferentes, embora os últimos possam levar aos primeiros. Outro aspecto marcante dos ruídos é que estes dizem respeito a variações não desejáveis. Da mesma maneira que os autores reconhecem que há vários julgamentos, como os artísticos, nos quais a variação é importante e desejável, preocupam-se com aquelas variações indesejáveis e que causam distorções tanto no plano individual – a mesma pessoa julgando o mesmo problema de forma distinta sem qualquer justificativa – como no plano coletivo – uma mesma instituição pública ou privada tendo diferentes julgamentos sobre os mesmos assuntos conforme a pessoa natural que fica a cargo de realizar o julgamento.

Vistos os ruídos sob esta perspectiva, percebe-se o quanto é legítima a preocupação dos autores, na medida em que identificam problema gravíssimo que pode comprometer os julgamentos individuais e, com maior razão, os julgamentos realizados no interior de organizações que dependem de várias pessoas. Daí ser fundamental não apenas estar atento para o problema do ruído, como também se pensar em possíveis soluções e estratégias para pelo menos mitigá-lo, por meio de técnicas englobadas no que os autores chamam de “higiene decisória”.

Como era de se esperar, a obra tem no Poder Judiciário um dos seus principais alvos de preocupação, mostrando o quanto pode ser injusto que as mesmas situações possam receber julgamentos tão distintos em razão das idiossincrasias ou do humor do juiz no momento da decisão ou mesmo de aspectos aleatórios que marcam o processo decisório, tais como o fato de o dia estar ensolarado ou chuvoso.

Entretanto, se o objetivo de redução do ruído é perfeitamente compreensível e necessário, a obra conclui no sentido de que ele é o principal, quando não o único objetivo dos julgamentos. Daí os autores sustentarem abertamente que o objetivo maior de qualquer julgamento é a acurácia.

Partindo também da premissa de que nossas mentes são instrumentos de mensuração, os autores descrevem julgamentos como

mensurações nas quais o instrumento é a mente humana. Consequentemente, um bom julgamento é um julgamento objetivo e com alto grau de acurácia²:

“Good decision making must be based on objective and accurate predictive judgments that are completely unaffected by hopes and fears, or by preferences and values.”

Em outras palavras, bons julgamentos são aqueles nos quais se neutraliza a própria humanidade do julgador. Daí por que, segundo os autores, regras, fórmulas e algoritmos normalmente vão ultrapassar os julgamentos humanos, pelo simples fato de serem *noise-free*³:

“In fact, many types of mechanical approaches, from almost laughably simple rules to the most sophisticated and impenetrable machine algorithms, can outperform human judgment. And one key reason for this outperformance—albeit not the only one—is that all mechanical approaches are noise-free.”

Não é sem razão que os autores defendem igualmente uma maior simplicidade dos processos decisórios, propondo que regras objetivas são normalmente preferíveis a standards ou normas abertas, exatamente porque reduzem a discricionariedade do julgador⁴:

“Rules have an important feature: *they reduce the role of judgment*. On that count, at least, judges (understood to include all those who apply rules) have less work to do. They follow the rules. For better or worse, they have far less room to maneuver.”

“Standards are altogether different. When standards are in place, judges have to do a lot of work to specify the meaning of open-ended terms.”

“By contrast, standards invite noise.”

“With rules, the costs of decisions are typically much lower.”

2 Op.cit., E-book.

3 Op.cit., E-book.

4 Op.cit., E-book.

Em relação à inteligência artificial, os autores são mais cautelosos, até porque não têm como negar o potencial de discriminação dos julgamentos daí decorrentes. Entretanto, advertem que há uma supervalorização do problema dos algoritmos, pois se ignora que os problemas dos julgamentos humanos podem ser ainda maiores⁵:

“Before drawing general conclusions about algorithms, however, we should remember that some algorithms are not only more accurate than human judges but also fairer.”

“The algorithm makes mistakes, of course. But if human judges make even more mistakes, whom should we trust?”

De toda sorte, os autores conferem especial atenção ao problema dos *noiseless, biased algorithms*, admitindo que a redução dos ruídos pode vir com um custo intolerável se houver aumento dos vieses preconceituosos⁶:

“The potentially high costs of noise reduction often come up in the context of algorithms, where there are growing objections to “algorithmic bias.” As we have seen, algorithms eliminate noise and often seem appealing for that reason. Indeed, much of this book might be taken as an argument for greater reliance on algorithms, simply because they are noiseless. But as we have also seen, noise reduction can come at an intolerable cost if greater reliance on algorithms increases discrimination on the basis of race and gender, or against members of disadvantaged groups.

Indeed, in this regard, algorithms could be worse: since they eliminate noise, they could be more *reliably* biased than human judges.”

Com isso, os próprios autores reconhecem as limitações dos julgamentos algorítmicos, sendo a questão da discriminação provavelmente o único exemplo em que os autores reconhecem que os algoritmos podem levar a resultados mais enviesados do que os julgamentos humanos.

Não obstante, há alguns outros pontos da obra que merecem uma maior reflexão. O primeiro deles, sem dúvida, diz respeito à própria premissa de

5 Op.cit., E-book.

6 Op.cit., E-book.

que os julgamentos humanos têm por principal ou único objetivo a acurácia. Ora, e a justiça? De que adianta a acurácia se o julgamento for claramente incorreto, preconceituoso, injusto ou inaceitável? O fato de haver igualdade formal e segurança jurídica em casos assim resolve? Adianta para as mulheres e outras minorias saberem que serão discriminados de forma igualitária, mantendo-se a precisão e o caráter *noise-free* dos julgamentos?

Ao salientar que a única preocupação dos julgamentos deve ser com acurácia, os autores acabam assumindo não só que existe um *tradeoff* entre acurácia e justiça, como que o referido *tradeoff* deve ser decidido incondicionalmente em favor da primeira. Entretanto, tanto a questão não é tão simples que, em vários momentos da obra, os autores parecem reconhecer o reducionismo da sua proposta, sugerindo alternativas mais abrangentes, em que propõem uma conciliação entre acurácia e justiça⁷:

“Undoubtedly, we need to draw attention to the costs of noiseless but biased algorithms, just as we need to consider the costs of noiseless but biased rules. The key question is whether we can design algorithms that do better than real-world human judges on a combination of criteria that matter: accuracy and noise reduction, and nondiscrimination and fairness.”

Não obstante, mesmo nessa “concessão”, a ideia é de que as discussões sobre justiça e valores sejam feitas também por algoritmos. Ao assim defenderem, os autores acabam menosprezando vários dos desafios que foram mapeados ao longo desta série, tais como as dificuldades de converter em fórmulas discussões mais complexas e sofisticadas sobre valores, o gravíssimo problema dos resultados não intencionais, dentre diversas outras.

Também não exploram os autores que a opção pelos algoritmos envolve normalmente uma delegação do processo decisório para a programação, sendo que os programadores estão sujeitos aos mesmos vieses e aos problemas aleatórios dos quais decorrem os ruídos. Conseqüentemente, a acurácia ou o caráter *noise-free* dos julgamentos algorítmicos pode vir às custas dos preconceitos, vieses e equívocos dos seus programadores, além de fatores

⁷ Op.cit., E-book.

aleatórios – os mesmos que levam aos ruídos - a que igualmente estão sujeitos no momento da programação.

Ainda que se trate de inteligência artificial, especialmente na modalidade *machine learning*, a transferência de julgamentos complexos sobre seres humanos para algoritmos implica claramente uma opção pela priorização das correlações e padrões encontrados nos dados – os quais podem simplesmente refletir e reproduzir preconceitos reais que existem na sociedade - em detrimento das necessárias reflexões a respeito de causalidade.

Não é sem razão que os autores desprezam o raciocínio causal e defendem abertamente a prevalência do raciocínio estatístico. Com isso, ignoram toda a riqueza das discussões sobre causalidade, inclusive a partir de diferentes *framings*, como foi abordado no artigo anterior, além de menosprezarem as diversas limitações das estatísticas, tema que foi também recorrente nessa série.

Também não é sem razão que os autores defendem julgamentos baseados em regras – e não em normas abertas ou standards – sem considerar que, especialmente no que diz respeito a decisões judiciais, o fracasso do positivismo legalista do século XIX decorreu precisamente da constatação de que, em uma sociedade cada vez mais complexa, é simplesmente impossível regular todas as situações por regras e que essa visão ingênua de segurança jurídica é obtida às custas de simplificações e reducionismos inaceitáveis.

Dessa maneira, se existem ruídos excessivos e inaceitáveis nas decisões judiciais e se os juízes estão exercendo indevidamente o seu poder decisório, a solução não é por fim a discricionariedade, mas sim contê-la a partir dos instrumentos adequados.

Aliás, é importante reiterar que, como os próprios autores reconhecem, divergências não são necessariamente ruins, ainda mais se forem bem colocadas, fundamentadas e consistentes, se resultarem de *framings* diversos e se contribuírem para uma visão mais ampla e diversificada do mesmo problema. Daí por que uma das dicas da obra, inspirada na “sabedoria das massas”, é precisamente a de obter o maior número de julgamentos possíveis a respeito do mesmo problema: “*Obtain independent judgments from multiple judges, then consider aggregating those judgments.*”

Nesse ponto, ainda que com pontos de partida totalmente diversos, a obra de Kahnemann, Sibony e Sunstein encontra-se com a obra de Cukier, Mayer-Shönberger e Véricourt⁸, explorada no artigo anterior, pois chega à conclusão de que a diversidade e o pluralismo de julgamentos podem ser extremamente importantes para o julgamento final. A partir desse *insight*, a própria ideia de segurança jurídica e formação de jurisprudência é ressignificada, uma vez que a uniformização, embora necessária a partir de um determinado momento, deve ocorrer apenas depois que houver tempo necessário para que a questão seja suficientemente debatida sob diversos ângulos, enfoques e enquadramentos.

Em outras palavras, até para se estabelecer a palavra final a respeito de um determinado assunto – e cogitar traduzi-lo em códigos e algoritmos - , há de se ter certo grau de maturidade no debate, objetivo para o qual uma divergência saudável e rica é fundamental. Buscar segurança, acurácia e uniformização a qualquer preço pode implicar a adoção prematura e incorreta de um ponto de vista a respeito de assunto que não foi suficientemente debatido.

además, considerando que assuntos humanos são extremamente dinâmicos, mesmo após a sedimentação de um determinado entendimento, há que se deixar válvulas de escape para atualizações, ressignificações e mesmo alterações. Tais raciocínios igualmente se aplicam a instituições privadas, que podem ter processos decisórios tão complexos quanto os que são travados no Poder Judiciário e igualmente precisam administrar a tensão entre diversidade, adequação, acurácia e coerência em seus julgamentos.

Dessa maneira, se não há dúvida de que redução de ruído e acurácia são objetivos que devem ser perseguidos, também não há dúvida de que estes não são os únicos propósitos das decisões, que precisam igualmente atender a pressupostos de adequação, justiça e ausência de discriminação. Por essas razões, não se pode optar jamais pela acurácia em detrimento da justiça e de princípios relevantes, como o da não-discriminação.

Aliás, os autores reconhecem que, na esfera preditiva, os modelos algorítmicos são consistentemente melhores do que pessoas, mas não muito melhores. Da mesma forma, reconhecem que a inteligência artificial tem

⁸ CUKIER, Kenneth; MAYER-SCHÖNBERGER, Viktor; VÉRICOURT, Francis de. *Framers. Human Advantage in an Age of Turmoil*. New York: Penguin Random House LLC, 2021.

melhores resultados do que modelos simplificados mas, mesmo assim, está muito longe da perfeição.

O simples fato de os julgamentos algorítmicos apresentarem tantos problemas e limites já é um importante aviso de que não é recomendável que sejam usados simplesmente para substituir julgamentos humanos. Pelo contrário, o mais prudente seria tentar obter uma postura de complementação entre os julgamentos algorítmicos e os humanos, a fim de que cada um possa compensar as debilidades do outro.

É por essas razões que, em que pese a legítima iniciativa de Kahneman, Sibony e Sunstein de diagnosticar e mitigar o problema dos ruídos, tal objetivo não pode ser pretexto nem para tornar os julgamentos algorítmicos os únicos nem – muito menos – para ignorar os benefícios dos julgamentos humanos, os quais já foram antecipados no artigo anterior e serão mais bem explorados no artigo seguinte, a partir da excelente obra de King e Kay.

Publicado em 15/09/2021

Link: <https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/discriminacao-algoritmica-e-o-problema-do-ruído-15092021>