



Por: Osvaldo Silva  
Professor Auxiliar do Departamento de Matemática e Estatística  
Faculdade de Ciências e Tecnologia da Universidade dos Açores  
osvaldo.dl.silva@uac.pt

# Arte ou Ciência dos Dados sem medo, mas com cautela!

Numa sociedade cada vez mais tecnológica, que se informatiza velozmente, tendemos a aceitar e a acreditar que com o recurso à tecnologia se consegue resolver a maioria dos nossos problemas de forma fácil e rápida, confiando que ela não é preconceituosa. O aparecimento ao longo das últimas décadas da designada “Ciência dos Dados”, interligada com uma forte componente tecnológica, permitiu o desenvolvimento de muitas ferramentas que ajudaram a superar as muitas tarefas rotineiras e levaram à necessidade de reduzir recursos humanos para as efetuar e tornar os processos mais eficientes. Ao processarem milhões de dados, numa fração de segundo, esses equipamentos tecnológicos conseguem indicar-nos informações úteis, tais como o percurso mais adequado para chegar a um determinado local, qual o hotel ideal numa determinada cidade para uma família com determinadas preferências ou requisitos. Esses equipamentos são ainda úteis ao nível das instituições privadas e públicas, contribuindo para a melhoria e eficiência dos seus sistemas e possibilitando a extração de conhecimento a partir dos dados a que têm acesso.

O cidadão comum ouve falar disto e, seduzido pela facilidade da tecnologia, tende a acreditar, de forma cega, em tudo aquilo que é transmitido como sendo uma verdade acima de qualquer suspeita. Para problemas simples do nosso dia a dia e em que há dados absolutamente mensuráveis constata-se que os resultados são fiáveis e que podem ajudar efetivamente, mas quando queremos resolver problemas de elevada complexidade deveremos ter algum cuidado no uso de recursos tecnológicos e os mesmos devem ser devidamente testados e não se cair na tentação de acreditar que estes resolvem todos os nossos problemas. Devemos ter a capacidade crítica e analítica e suspeitar sempre de tudo, com imparcialidade, ponderação e capacidade de reconhecer o erro. Os cidadãos devem ser exigentes e vigilantes e não cair na tentação de somente idolatrar a ideologia tecnológica. Só com um confronto sério, entre os resultados obtidos com recurso à ciência dos dados (data mining, big data, ...), que recorre a algoritmos e a técnicas estatísticas, e a realidade que vivenciamos diariamente, é que se pode avançar como sociedade, de forma segura e refletida. Numa sociedade em que algumas empresas ligadas a esse setor são excessivamente ambiciosas e perspetivam obter o lucro rápido, estas tentam vender essas soluções (programas informáticos) como se fossem o remédio milagroso para a eficiência das organizações. Por isso, os cidadãos, em geral, e em particular os gestores das organizações, devem ter o cuidado de se informarem adequadamente sobre o funcionamento dessas soluções existentes, para que possam agir de forma competente, sem caírem na tentação do facilitismo. Por vezes, só quando já é tarde demais é que sentem que foram enganados, o que pode acarretar elevados custos para as organizações e para os seus colaboradores. Atenção que uma grande parte talvez nunca se sinta enganada porque nunca pensou nisso. Em geral, os cidadãos e as instituições desconhecem os verdadeiros contornos de funcionamento dessas ferramentas, muito badaladas na atualidade,

que recorrem à inteligência artificial. Exigir informação é um ato de cidadania na defesa dos nossos direitos e liberdades a nível individual e em prol da sociedade em que estamos inseridos. Toda essa atividade de recolha, armazenamento e análise de dados, alguns dos quais referentes a registos sobre todos nós, deve ser transparente e devidamente regulada, porque estamos ao mesmo tempo fazendo parte dessa atividade, muitas vezes se calhar de forma inconsciente, como produtores, consumidores (clientes) e até como potenciais alvos do sistema em nome da eficiência.

Nos tempos que correm, podemos recolher dados sobre as pessoas, extrapolá-los e até manipulá-los com um certo à-vontade, criar ferramentas analíticas que podem afetar as nossas vidas, para o bem e para o mal, sem que existam por vezes mecanismos de proteção que salvaguardem os direitos de autor e, ao mesmo tempo, garantam que os benefícios para a sociedade superem os aspetos negativos associados aos mesmos.

Com o manancial de ferramentas utilizadas, com recurso a variadas técnicas estatísticas e com o suporte de conhecimentos de informática, não é plausível que o cidadão comum tenha capacidade para decifrar se os resultados apresentados podem ser considerados como válidos ou não para a realidade em estudo. Embora possa existir um código de conduta para os especialistas em dados, deverá ser o Estado, através de um organismo autónomo e independente, a garantir a idoneidade dessas ferramentas (algoritmos) desenvolvidas e que estas funcionem sem causarem prejuízos aos cidadãos. Muitas vezes a adoção progressiva destas ferramentas, por parte das organizações, não é só por razões da eficácia da tecnologia, mas também devido à contenção de custos.

Com o fascínio pela tecnologia e em nome do acesso rápido à informação, de modo a uma gestão mais eficiente dos seus recursos, os países e as suas organizações tendem, cada vez mais, a desenvolver bases de dados com informação acerca dos seus cidadãos. Atenção que as bases de dados são feitas por pessoas que podem colocar informações que podem constituir fator(es) de discriminação, que podem induzir a que haja enviesamentos aquando da análise desses resultados e das conclusões retiradas a partir dos mesmos.

Um dos problemas situa-se ao nível da área da saúde, como sucede por exemplo nos EUA, com base num estudo publicado em outubro de 2019 na revista Science, onde é referido que um algoritmo utilizado nos hospitais norte-americanos beneficia menos os negros do que os brancos nos programas criados para melhorar o atendimento de doentes com necessidades médicas mais complexas. O problema não é somente o algoritmo utilizado, mas toda a abordagem do sistema é problemática. Não é fácil corrigir o enviesamento do algoritmo, seja na área da saúde ou na área da justiça, por exemplo. Desenvolver uma ferramenta 100% imparcial é quase impossível porque potenciais enviesamentos poderão também ocorrer ao nível da recolha de dados.

Na área da segurança e da justiça são também cada vez mais utilizados programas informáticos que prevêm o comportamento humano, para pres-



supor, por exemplo, os locais onde serão cometidas infrações e em que residências é mais provável existir violência familiar. Muitas vezes são utilizadas técnicas de aprendizagem automática, com recurso a múltiplas árvores de decisões (compostas por questões pertinentes a serem analisadas) até se obter uma previsão.

Na área social, para a avaliação de problemas sociais, existem várias experiências com a utilização de algoritmos, com o intuito de preverem as necessidades sociais, como sejam, por exemplo, em alguns municípios do Reino Unido. Conjugando informações acerca dos seus habitantes, obtidas por diferentes instituições (e.g., Polícia, Serviço Nacional de Saúde, Instituto de Emprego e Segurança Social e do Município), os utilizadores dessas plataformas têm acesso a registos sobre emprego, problemas com álcool, drogas, gravidez precoce e violência doméstica, entre muitas outras. Cerca de 170 mil habitantes de Bristol (no Sudoeste de Inglaterra) são seguidos por uma ferramenta informática, tendo cada pessoa recebido uma pontuação de 1 a 100, indicando a possibilidade da prática de atos ilegais, de saída do município, e de praticar maus tratos a crianças, entre muitas outras informações. Numa questão de segundos, o sistema consegue identificar o risco associado a um determinado cidadão dessa cidade, enquanto os serviços sociais demorariam muito tempo, muitas vezes anos, a detetá-lo. O uso desse sistema tem também o intuito de detetar e reduzir a ocorrência de fraudes ou de antecipadamente prevenir problemas sociais que sejam muito onerosos. Por outro lado, com base nas informações comunicadas pelo sistema podem ser desenvolvidas estratégias atempadas no apoio social para minimizar fatores de risco, como sejam o abandono escolar, a violência doméstica e a delinquência.

Contudo, a experiência muito recente, com o aumento de empresas a atuarem nesse setor e com sistemas pouco credíveis a darem resultados desfados da realidade, já levou a que alguns municípios (por exemplo, Sunderland, no Noroeste de Inglaterra) recusassem na utilização desses algoritmos e não renovassem contratos com as

empresas privadas que prestaram esses serviços. Um dos motivos para essa renúncia foi o facto de o algoritmo ter detetado erradamente habitantes dessa cidade como pessoas potencialmente fraudulentas, o que resultou num atraso significativo no pagamento de subsídios de assistência social.

Atenção que o sistema pode ajudar, mas não substitui o julgamento de um profissional e as decisões devem ser sempre bem analisadas e devidamente ponderadas. O sistema é apenas um fornecedor de informações que estavam dispersas em diferentes serviços. É preciso estarmos conscientes em relação ao que o sistema nos diz, mas somente com o auxílio de pessoas com competências e bem-intencionadas é que essa tecnologia poderá ser utilizada em benefício de todos nós.

Muitas vezes os criadores de algoritmos, como sejam os matemáticos, engenheiros e programadores informáticos, não fazem a devida interconexão entre os dados e os algoritmos que concebem, e na maioria dos casos não há uma devida reflexão ética e crítica sobre o que desenvolverem e as potenciais implicações positivas e negativas que as mesmas poderão ter nos seus concidadãos. Cada um faz o seu trabalho de forma parcelar e não há comunicação entre os seus criadores e os utilizadores. Os algoritmos nunca serão perfeitos, mas podem ser continuamente melhorados, de forma a serem minimizados eventuais erros. O que é imprescindível é abandonar o hábito que temos de confiar completamente nas máquinas. Não se esqueça de que estas foram feitas pelos Homens, por isso as máquinas padecem das falhas humanas. Isto é os algoritmos são concebidos para terem em conta a nossa falibilidade enquanto seres humanos.

Só com as pessoas bem informadas acerca da utilização dessas ferramentas, dos seus benefícios e riscos associados, dos cuidados a ter na análise dos resultados obtidos e das conclusões retiradas é que poderemos avançar sem medos, mas de forma cautelosa. Atenção que já estamos a viver num mundo em que algumas das principais decisões são tomadas pelas máquinas, sem que nós tenhamos consciência disso. Já refletiu um pouco em relação à forma como as suas decisões são tomadas?