

1) Big Data diz respeito à acumulação de dados, tanto desestruturados quanto estruturados, que podem ser coletados de inúmeras formas, como, por exemplo, sensores, aplicativos, lojas em geral, sites de busca, de venda, de relacionamento.... Há certa facilidade em se analisar dados estruturados, por outro lado, dados desestruturados requerem formas mais engenhosas para sua avaliação. Entretanto, ambos os dados podem ser analisados, preferencialmente por uma máquina (dado grande capacidade), a fim de gerar-se informação.

O dado coletado tem seu valor porém é apenas a partir do seu processamento que a informação é gerada. O valor agregado na informação produzida, que foi coletada a partir de incalculáveis fontes e de imensuráveis formas distintas, é surpreendente. Visto que, a partir dessas informações, decisões podem ser tomadas com base em verdadeiras análises, gerando resultados muito mais favoráveis, daí seu valor e importância.

2) “O Big Data oferece significado, correlação, mas não oferece a causa”. Um dos exemplos apresentados no texto a partir dessa frase é sobre a voz de Morgan Freeman. Pergunto-me se a causa seria de suma importância para a tomar-se uma decisão nesse caso. Por exemplo, já que tal informação pôde ser concluída, por que não fazer mais narrações de cunho intelectual com o ator citado? Já que pessoas com alto QI provavelmente prefeririam escutar algo desse gênero com a voz do ator. Certamente que seria mais valioso saber a causa, já que isso seria uma informação suplementar, porém não creio que esse tipo de incerteza seja capaz de eliminar por completo a tomada de decisão. Como será visto no item 3, isso pode gerar discussões éticas.

“A análise do oceano de informações pode revelar um padrão, uma correlação, um significado antes oculto e, quase sempre, ajuda a prever o futuro”. A palavra “oculto” inserida nessa frase é intrigante. O Big Data, unido com inteligência artificial e aprendizado de máquina podem gerar resultados que para nós humanos passam despercebidos. Há um exemplo na área de reconhecimento de imagem, em que foi fornecido grande quantidade de dados, neste caso, imagens de câncer, para um algoritmo adaptado à inteligência artificial e aprendizado de máquina. Ele foi capaz de, não apenas identificar o câncer, mas aumentar de 9 para 12 as características até então conhecidas ao interpretar a imagem de câncer. Ou seja, o algoritmo percebeu mais 3 feições nas imagens que antes passavam despercebidas nas observações dos médicos.

3) A grande vantagem é o embasamento que a Big Data fornece nas tomadas de decisões. Humanos não são capazes, por si só, de analisar imensas quantidades de dados. Porém, com o auxílio oferecido por máquinas os dados podem ser processados, tornando-se informação. E essa informação oferece mais sustentação em decisões. Com isso se pode citar inúmeras vantagens.

Um primeiro exemplo seria o (já citado acima) processamento de imagens. Isto é, fornecer vastas quantidades de dados e utilizar o aprendizado de máquina e inteligência artificial (e possivelmente aprendizagem profunda) para ensinar a máquina a reconhecer imagens. O reconhecimento de imagem tem outra extensa gama de vantagens, por exemplo a identificação de câncer, ou qualquer outro evento que se queira identificar a partir de imagem.

Outro exemplo poderia ser a famosa inteligência artificial que recomenda filmes, produtos, livros, podcasts ou outros produtos (dependendo da IA). Essa IA necessita ser treinada, e seu treino consiste da utilização dos dados dos comportamento dos consumidores, ou

seja, Big Data. Ela é capaz de analisar os dados coletados e recomendar algo com maior probabilidade de satisfazer a necessidade de seus clientes.

Por outro lado, um dos pontos negativos pode ser o de discriminação. Suponha-se que um sistema de contratação de pessoas tenha sido treinado através de Big Data e aprendizagem de máquina, e instruído a contratar pessoas que tivessem o mesmo perfil dos funcionários de alto desempenho da empresa. Parece uma boa ideia, mas, e se tal algoritmo estivesse (não sabemos os parâmetros utilizados pelo sistema) eliminando pessoas com alta probabilidade de depressão futura? Com alta probabilidade de engravidar? Ou ainda, contratando pessoas agressivas, que podem ser bons funcionários e gerarem lucros, mas possuem um histórico violento? Apesar de tal algoritmo realmente contratar bons funcionários, será isso ético? Eliminar pessoas baseadas em inferência? Voltamos ao ponto do número 2, em que temos a informação útil, porém não sabemos a causa específica da decisão.

Outro ponto negativo é possivelmente o tipo de dado que é fornecido, por exemplo, para uma AGI, isto é, inteligência artificial geral, capaz do processo autorretroativo de aprendizagem. Como visto no último trabalho sobre as leis de Asimov, nota-se que o tipo de dado, ou informação, entregue a uma AGI é importante. Caso sua base de informação for a internet como um todo, isso pode torná-la muito parecida com um ser humano. Há vantagens nisso, porém há também imensas desvantagens, visto que humanos não são totalmente capazes de possuir um bom relacionamento com o ambiente em que vivem (incluindo relacionamento com outros seres humanos), e isso não deveria ser ensinado a uma máquina. Portanto, a qualidade do dado fornecido a uma máquina possui relevância, e ao tratar-se da criação de uma possível futura AGI, algo catastróficamente imperfeito pode ser criado. Vide o chatbot criado pela Microsoft, que a partir de dados fornecidos pelos usuários do Twitter (humanos), tornou-se “um ser” de péssimo caráter.