

Research World: January, 2020

Bias is natural. It's also dangerous

Just a few weeks ago, Apple announced with great fanfare its new credit card. Behind the scenes was JP Morgan Chase, entering the credit card market for the first time. Very soon it emerged that men were receiving 10 to 20 times more in credit lines than their wives who were giving identical information to the bank in terms of their financial status. The only difference in their applications was their gender. Bank officials were at a loss to explain the discrepancies and blamed the algorithm used to determine credit-worthiness.

Therein lay the problem. The algorithm had somehow incorporated biases that favored males as credit risks over females. Machine Learning then very likely had built on the algorithm and magnified the bias.

The fact is that algorithms are written by humans and whatever biases these humans may have (conscious or unconscious) can very easily end up being incorporated into the algorithms themselves. And, in an age where we seem all too willing to trust the black boxes of Machine Learning, such biases become embedded, exaggerated and, in some cases, permanent. This is bad enough when the bias in question is unconscious, but there have been cases recently where data scientists have proudly boasted of having deliberately introduced biases into their algorithms because this “reflects more accurately the world we live in”. Such an argument is not only specious, it is downright dangerous. Imagine that AI and ML had existed in the age when eugenics was all the rage and was used to explain away racism. Had those biases (or beliefs) been incorporated into the algorithms of the time, it is entirely possible that the US Civil Rights movement would never have happened.

It's up to us

If we recognize that biases are inherent in algorithms, what can we do to mitigate the danger? First off, we need to elevate the conversation about bias, as well as about integrity and ethics in all that we as insights professionals do. The world's major trade and professional insights associations are all putting a major effort into doing this, raising the subject not only within the industry itself but also with governments and academia. But the conversation needs to occur earlier in the process of producing insights professionals and data scientists. Awareness of the issue of bias needs to be introduced at all levels of tertiary education and across a much broader swathe of subjects than merely business analytics or market research.

While we are on the subject of market research, let us also acknowledge not only that bias exists at many levels within the research process but that, in many instances, this inconvenient

truth has been forgotten or is willfully ignored. While seasoned researchers (especially in the political and social fields) are fully aware of a myriad of biases – explicit, implicit, mode, question order, question wording (you name it) – and work hard to mitigate it, many others may only be dimly aware of what they are or even that they exist. In an age of convenience samples, digital sample sourcing, automated scripting and DIY platforms, discussion of bias has all but fallen away. Yes, those who have earned degrees in MR or who have taken training courses are very likely to be aware, and may even diligently apply mitigation, but they are a small minority of all of those who are now practicing some form of primary research. Which means that the good old maxim “garbage in, garbage out” is more relevant than ever, and that there is an omnipresent danger of decisions being taken on the basis of bad data. For this to be the case at all – and for the industry not to deal with it in an age that combines data overload with data illiteracy – would not only be unfortunate. It would be unethical.

A call to arms

It is time for us all to have a very robust conversation about the issues of bias, integrity and ethics in the insights and data sciences space. Convenience should never prevail over data quality (and please don't quote “directional” or “good enough, move on” at me). However we derive our insights – whether through primary research or data analytics or a combination of both – we need a sustained campaign of education to introduce these concepts to a much larger audience than is happening today. To my mind, all BBA and MBA students should be taught the basics of bias and how to mitigate for it. Only in this way can we ensure that we remain true to our roots – impacting business success, guiding social policy and political decision-making, all based on data of the highest quality and integrity.