Why have evaluators been slow to adopt big data analytics?

Description

Tweet

This is a question posed by Michael Bamberger in his blog posting on the MERL Tech website, titled Building bridges between evaluators and big data analysts. There he puts forward eight reasons (4 main ones and 4 subsidiary points). None of which I disagree with. But I have my own perspective on the same question and posted the following points as a Comment underneath his blog posting.

My take on â??Why have evaluators been slow to adopt big data analytics?â?•

1. â??Big data? I am having enough trouble finding any useful data! How to analyse big data is â??a problem we would like to haveâ??â?• This is what I suspect many evaluators are thinking.

2. â??Data mining is BADâ?• â?? because data mining is seen as by evaluators something that is ad hoc and non-transparent. Whereas the best data mining practices are systematic and transparent.

3. â??Correlation does not mean causationâ?• â?? many evaluators have not updated this formulation to the more useful â??Association is a necessary but insufficient basis for a strong causal claimâ?•

4. Evaluators focus on explanatory models and do not give much attention to the uses of predictive models, but both are useful in the real world, including the combination of both. Some predictive models can become explanatory models, through follow-up within-case investigations.

5. Lack of appreciation of the limits of manual hypothesis formulation and testing (useful as it can be) as a means of accumulating knowledge. In a project with four outputs and four outcomes there can be 16 different individual causal links between outputs and outcomes, but 2 to the power of 16 possible combinations of these causal links. Thatâ??s a lot of theories to choose from (65,536). In this context, search algorithms can be very useful.

6. Lack of knowledge and confidence in the use of machine learning software. There is still work to be done to make this software more user friendly. Rapid Miner, BigML, and EvalC3 are heading in the right direction.

7. Most evaluators probably donâ??t know that you can use the above software on small data sets. They donâ??t only work with large data sets. Yesterday I was using EvalC3 with a data set describing 25 cases only.

8. The difficulty of understanding some machine learning findings. Decision tree models (one means of machine learning) are eminently readable, but few can explain the internal logic of specific prediction models generated by artificial neural networks (another means of machine learning, often used for classification of images). Lack of explainability presents a major problem for public accountability. Public accountability for the behavior and use of algorithms is shaping up to be a BIG issue, as highlighted in this weekâ??s Economist Leader article on advances in facial recognition software: What machines can tell from your face

Update: 2017 09 19: See Michael Bambergerâ??s response to my comments above in the Comment section below. They are copied from his original response posted here http://merltech.org/building-bridges-between-evaluators-and-big-data-analysts/

Category

1. Uncategorized

Date
13/06/2026
Date Created
09/09/2017
Author
admin