Vicious Delicious Bias: the relation between humans, computers and truth.

Vincent de Wit (6970877)
v.j.dewit@students.uu.nl
MSc Artificial Intelligence
Utrecht University

April 10, 2020

Contents

1 Introduction

Even though bias is clearly a bad thing when making models of reality, bias is also responsible for good things. For example we as people have conceptions of how the world works or what other people think or what we should be doing. This can result in people becoming friends or creating a good working business. Of course such ideas can also have bad consequences like fight between people due to misunderstanding or it could results in strong beliefs that are false. But it can be said that many behaviours we exhibit stem from our conception of what is true. Even though our model of reality is not completely right it allows us to act in the world. When our model of reality is biassed it means there is information that we are unaware of that would result in different model, which consequently could very likely change our behaviour. Even though this bias can have harmless effects, a model should be changed when it is understood that it misses some information. To make matters worse it can happen that either we are biased to disregard the new information or we are biased to accept the new information too willingly. This interplay, between a model and reality, also plays in science. Although in science the notion of a model and the process of information gathering is formalized. Nevertheless I will argue that models and bias are inseparable even/especially in science.

It would have been fun to spend this paper indulging in all kinds of philosophers that I find interesting like John Stuart Mill, Witgenstein and Derida. But instead this essay will be mainly focussed on what a model is, methods used in models and choosing a model. All these things but mostly models themselves are connected to bias. To indulge further we could have spend this essay discussing sceptics like Phyro and Montaigne. Phyro for example made a great analogy of blindness with biassed models. And Montaigne has created a wonderful method: suspension of judgement, which we could use to deal with the inaccuracy induced by bias in models. So instead of revolting like Sextus Empircus this essay will discuss three papers written by three different people. These people are, in the order by which I will discuss their paper, Jon Williamson, Richard J. C. M. Starmans and Wolfgang Pietsch. The first paper will be discussed in light of hypothesis choice and model selection which are respectively related to human and computer activities. Then epistemology and research methodology will be discussed in relation to the second paper which are about truth itself en humans methodology to find this so called truth. And the last paper will be discussed by means of data-intensive modelling methods which are executed by computers and their relation to causality which humans can understand but computers notoriously have trouble with.

2 Hypothesis choice and model selection

The process of hypothesis choice in science and model selection in machine learning are both forms of systematizing [3]. This term means finding generalities on the basis of evidence. Williamson explains that these concepts besides sharing a category have some things in common. This mostly regards the role it takes with regard to the subject of a model.

Hypothesis choice is guided by a scientists beliefs and understanding of his field. Furthermore the choice of hypothesis might be based on whether data can be gathered on the specific subject. The hypothesis plays a big role when making a model as it determines in an abstract way what the model is about. Also the way the hypothesis is structured can bring potential bias into light, due to negative and positive formulations.

When a scientist makes a computer model it is important to make a balanced choice which model eventually to use. As the scientist can choose to train an architecture multiple times on the data set. Each model configuration can have slight but also larger differences. And since eventually one model configuration should be chosen good heuristics should be used to help in the decision. The scientist can't get around being biased even if very little therefore these heuristics can be very useful. Though a good heuristic should be as objective and indicative possible while minimizing possible bias. More importantly machine learning models are well known to capitalize on bias in data to improve their accuracy ratings. Therefore when choosing a computer model it will be mostly important to focus on mitigating such type of bias.

So far there are two bias factors in this process namely the bias in regards to choosing the model configuration from a set of configurations and the data the model trains on. Next to this the hypothesis choice also showed a similar structure where the scientist based on his knowledge, which could contain bias, and personal interest, which is by definition biassed. Are used to form and choose a hypothesis to research.

The intersection of these fields is automated scientific discovery. This would reduce the bias that is added by humans quite significantly. Since in a non automated version people still need to make decisions in regards to the model and the hypothesis. In this automated version this bias factor is factored out. Though this reduces the human bias significantly there is of course still bias transferred from humans through the architecture of the automated system.

3 Epistemology and research methodology

The study of the truth has advanced along two lines namely epistemology and research methodology [2]. In epistemology the nature of truth is studied while research methodology is used to do empirical research. There is a clear distinction between these two as epistemology is concerned with what is true or real. While methodology is concerned with creating trustworthy methods to derive what is truth based on measurements. To make this distinction clearer epistemology could be characterized as more personal while methodology is less personal as it is based on interpretation of measurements. This seems to be the case as people often have intuitive feelings about what is truth compared to statistical representations of measurements.

We humans have an intimate relation with truth, the conception of truth each of us cherishes can be said to be fragile. Those conceptions include and are not limited to hypothesizing, thinking, understanding and deducing. The fragility of this truth lies with the interests of people themselves and those of other people. These interests consequently alter and/or influence the beliefs of people towards a different conception of truth. The notion of truth a person holds is very important for that person personally. As what each person believes plays a big role in who they think they are, how to behave but also what to about other topics. This propagates on towards the larger scale of many people showing that many structures of power are in the business of truth. Structures than can be thought of are religions, science but also countries. As there are many interests implicated with what people believe it is accompanied by bias. This clearly shows that a finding what is actually truth is important and so that epistemology can be very useful.

To make our notion of truth less susceptible to ill influences we started using statistics. This would help us in finding out what was true and what was not by means of calculations and tests. Even though this sounds like a great idea the relation between statistics and truth is rather problematic. Of course statistics solves many issues but it can also easily help with misinformation. Either due to numbers that have been tempered with or by indiscriminately giving averages, standard deviations and other numerical abstractions. For example giving a percentage can be useful though when a low percentage is easily obtainable due to a high total amount and a minimal already being a problem, a small percentage can already be noteworthy though it can easily be under valued. Such translation from measurements to statistical values to appreciations of the values are liable to bias in each translation gap. Summing such bias can in the extreme case lead to a different story than the measurements tell.

This begs the question whether the concept model is eroding. Since models have become more abstract mathematical structures that are mostly focused at making the correct predictions. Instead of models being also focussed at explaining the mechanism that creates the measurements. This absence of explanation gives interpretation space. Due to the fragility of truth and the difficulties surrounding correctly interpreting statistics we can speak of the bias surrounding models as erosion.

4 Data-intensive modelling and causality

We tend to believe that data-intensive modelling has the short coming of not being able to handle causality. But as Pietsch shows, data-intensive modelling also known as horizontal modelling is able to identify causal relevance much like classical hierarchical methods [1]. The reason for the distinction of hierarchical and horizontal is based on the way relation are structured. In hierarchical methods a hierarchy of relations is systematized often by use of explainable formula's. While in horizontal methods there is no hierarchy of formula's instead all data is considered in the same horizontal "class level".

The data-intensive or horizontal modelling methods utilize correlation to locate causal relations. But we all well know that correlation is not enough to identify causation. Though this maybe true the horizontal methods therefore make use of eliminative induction methods. These methods are variations of a simple idea that is also common in sudoku's namely induction. By cross linking the correlation with the circumstances; causation can be identified based on correlation. This suggests an account of causation based on difference-making and is also known as the difference-making account. By understanding the correlation at play certain correlation can be excluded or eliminated based on circumstances. By repeating the procedure a complete understanding of the correlations under all circumstances can be created which induces a causal account of the relation.

Eliminative induction methods:

- method of exclusion
- method of elimination
- method of difference
- method of agreement

The standard method can be said to be hierarchical where the data-intensive method is horizontal. These standard methods are mostly parametric models that use distribution functions to describe the data by fitting their parameters. These models are grounded by theorizing on the data which results in an explainable model. When theorizing about the data a distribution function can be deduced after which only the parameters like the mean and the standard deviation need to be chosen in order to derive a model that can make predictions. It is important to theorize about the data as only then can the choice for distribution function be grounded in an explanation. This eventually results in more reliable predictions since the predictions are based on a distribution which corresponds to the theorized causal relations.

As the non-parametric models do not use parametrized functions to model the data, they find as described the relations based on cross linking correlations. This also means that where the models used to be based on an explanation of the causal relation they do not with these horizontal models. This could be called, knowledge without understanding. This could mean that since we understand less of the model it is easier to form bias in interpreting the model, since it is not grounded in an explanation any more. Also due the model cross linking all the correlations a lot of trust has to be put in the model and with that potential biassed data is more difficult to detect.

5 Conclusion

Hopefully this convinces you that models involve bias towards what truth is, what to model, how to model, what modelling can and can not do, but also the predictive ability of models in relation to their limit in scope. This shows that bias is inseparable from models. Even though bias should be minimized and hopefully this is also understood from the examples bias is inseparable from models. Therefore bias is also part of truth or reality and sometimes should therefore be even part of models. Completely removing bias therefore is not only impossible but also unnecessary. If models are made bias has to be a factor to consider while creating the model. As bias can creep into many aspects of the model. If it is not the data it might be the approach or even the model itself. Nevertheless a model with bias will often be better than no model at all. Therefore sometimes the trade-off for bias or error in one place might be worth more precision in another. Furthermore interpretation seems to be a big root of prejudice and bias. As this contains a translation step from abstract to story. This means that people that do not know the numbers should be careful with accepting any story. And those that know the numbers should be aware of people that make stories from them. And even though whatever model of reality we have as biassed as it may be we should still talk about it and share the views. It could be that the bias in our view remains and that it may even spread as a common conception which would be quite bad. Though talking and sharing the perspectives given that those are inseparable from bias also are the only way to allow us to distinguish bias from truth. And so is an important tool to eventually reduce biassed views. A good example for this are the horizontal modelling methods, before knowing about these methods I believed that correlation is not enough for causation. And even though this is still close to true the horizontal methods are able to come very close to showing causation. As long as I had not known about these methods I would of course had believed that it was likely not unwise to believe that correlation is not enough for causation. This also shows that it is easy to be convinced of something in such a way that the opposite first needs to be proven in order to change the believe. This would mean that we would need to either falsify and verify really many hypotheses or that we need to be able to create models given that bias remains in the model. Such that the model can do what it is made for and is understood to be made for that what it is supposedly made for.

Indulgement

To continue the indulgement from the introduction by elaborating on the explained ideas of Montaigne and Phyro we should first state something about scepticism. These two philosophers are namely known as sceptics. Scepticism is about doubting towards dogma, by doubting dogma's and presumed truth one is taking sceptic stance. Phyro is known as the first person who was a sceptic. He was sceptical towards the reality he perceived and therefore acted by disregarding what his senses notified him of. When there would be a cliff that he had seen he would not believe that that cliff actually existed and would walk as if the cliff was not there. This resulted in his friends having to keep him from falling of this cliff. There is here a very strong analogy the discussed concepts. Namely as we form a model of reality using our sense how can we now it is true.

Well we can not this model is by definition biassed but Phyro showed that even though he might be right in that the reality his perception is not unbiased it is able to keep him from falling of a cliff, if he would trust his senses enough that is. Montaigne takes a different approach as a sceptic. He is famous for at least two concepts namely essayer and suspending judgement. Where essayer is all about making an effort to try something which is useful because through such a process one is able to test their model. Suspension of judgement goes very well with this namely by suspending judgement one can let things move while observing what happens without colouring the events by judgements. This way the concept of essayer can be strengthened by presumed great unbiasedness toward the events resulting in a more thorough test of the model. As the model can be brought into relation with a conception of the events supposedly separate from the models interpretation of those events. This way a separate conception of events is conceived next to the model interpretation creating relativity on basis of which the model could be improved. To end in terms of Sextus Empiricus; Against Biassed Biassedness!

References

[1] Wolfgang Pietsch. The causal nature of modeling with big data. *Philosophy and Technology*, 29(2):137–171, 2016.

[2] Richard J. C. M. Starmans. *The Predicament of Truth: On Statistics, Causality, Physics, and the Philosophy of Science*, pages 561–584. Springer International Publishing, Cham, 2018.

[3] Jon Williamson. *The Philosophy of Science and its relation to Machine Learning*, pages 77–89. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.