Programming in Finance II - Weapons of Math Destruction

Matteo Bottacini

May 12, 2020

"Big Data processes codify the past. They do not invent the future. Doing that requires moral imagination, and that's something only humans can provide. We have to explicitly embed better values into our algorithms, creating Big Data models that follow our ethical lead. Sometimes that will mean putting fairness ahead of profit."

Cathy O'Neil

Weapons of math destruction (WMD) are mathematical models or algorithms that seek to provide a representation of real-world processes like the teacher quality, the recidivism risk and the creditworthiness. But having wrong or harmful outcomes often reinforces inequality, tending to punish the poor and the oppressed of society instead of making the rich wealthier.

Models are abstract representations of real-world situations, designed to produce a simplified version of reality. They are about translating a set of hypotheses about the behavior of predetermined features and labels into numerical predictions with the idea that everything could be empirically explained, eliminating human prejudice, misunderstanding, and bias into software systems that increasingly manage people's lives.

But what about the subjectivity of the modeler? Wrong assumptions lead to wrong results.

According to the author WMDs have three things in common: opacity, scale, and damage.

Opacity prevents the possibility of understanding how their specific output has been generated. Algorithms may contain miscodes, and opacity keeps such miscodes hidden; as a result, algorithmic opacity impacts on a fair result. On the other hand, transparency could lead participants involved in the model to handle those inputs affecting results to achieve a better performance. Thus, risking that not the model fits reality, but reality adapts itself to the model.

Scale is about the ability of the model to generalize well even considering multiple participants who weren't involved in the training procedure. Models improve their results based on what, and how, they have learned. Dynamic models keep updating and adjusting their learning procedure in order to evolve as the real world does. Feedbacks are crucial to understand if the model assesses performance saying something about reality or if it is off-track. With no feedbacks, no adjustments and no developing, there is both the risk of overfitting and underfitting. The former occurs when the model performs on the training data but is unable to generalize on unseen data. The latter turns up when it poorly performs even on the training set; so that is unable to acquire a clear relation between features and target.

Damage occurs when WMDs affect people on critical life moments, considering those victims as collateral damage of the model. Thanks to algorithms and Big Data analysis many people found the ideal job on LinkedIn or the love of their life on Match.com. Due to the same software other people cannot assess a predetermined job position or have access to credit lines because they are considered too risky.

This kind of problem might occur when models are based on proxies rather than correlations between variables that are not considered yet. Proxies are easier to manipulate since they simplify the complexity of the reality they represent. Sometimes models tend to over-simplify real-world situations that lack in generalization as they are not good enough to let people take the right decision.

Furthermore, some models try to say something about the future assuming that it will resemble the past, forecasting past verified patterns thinking the past will repeat itself again.

Models can be powerful tools that help humans take decisions. Mathematics and algorithms should be a fundamental support in people's decisions, but not the unique tools. Sometimes models cannot capture human irrationality in the same way as reality is not always following a precise scheme.

Big data analysis is driving companies' decisions affecting people's lives. Understanding how to implement it and the manner in which it is done will be a key factor in an increasingly competitive world, both if one is a modeler and if one is a contestant.

"Humans will always be involved in creating, guiding, and learning from the algorithms that assist models. It is not human decisions that are automated, but rather, the process of generating the information that drives those decisions".