Name: Leonardo Vida
ID: 6557929
Book: Nate Silver, The Signal and The Noise: why Some Predictions Fail - but Some Don't

How can you predict where the next terrorist attack is going to hit? What about the next earthquake? What if I would ask you to bet some of your money on a game of *Texas hold 'em*, or bet on a baseball game, or ask you to decide whether to go long or short on the financial market with your savings, how would you start approaching these problems? Which data would you use? How confident would you be in your bet?

Nate Silver, the author of "*The Signal and The Noise: why Some Predictions Fail - but Some Don't*" (2012), believes that the prediction and the associated bets of the majority of us on the previous questions would likely result in failure. Silver's book sets out to evaluate predictions in a variety of fields and, learning about failures, plans to become a skeptic's manual of best practices to adopt when approaching prediction problems that deal with big data.

"Why would my prediction fail?" you ask? Silver says that there are things we can predict, and we are good at predicting, and there are many other things that cannot be predicted accurately and, as expected, when predicted are predicted wrongly. Originally the failure of these predictions arises because the overwhelming quantity of noise with respect to signal in our data leads us to see patterns where none exist. However, throughout the book, Silver makes other examples to point out various other mistakes we do while formulating predictions. Here are some of the most interesting ones. First, perverse incentives and inability to take into account model uncertainty. These are incentives such as not to short-sell stocks too much, because of both cultural and legal reasons, and the inability to fully understand the risk, as it happened with default rate of mortgage securities during the housing bubble of 2008. Secondly, overconfidence in the environment of the prediction. In a chaotic system such as weather or earthquakes predictions, a small change in the input translates into a large difference in the output and this leads to overfitting existing data. This is the main reason why weather forecasts are still not able to reliably predict more than a week in advance and we are still unable to predict where and when the next earthquake is going to hit. Thirdly, extrapolation of the trend. Here failure of the model derives because of the assumption that the current trend will continue indefinitely. Examples of this failure are numerous flu and HIV campaigns based on the wrong prediction of the spread of the disease.

Summarizing, poor predictors often share the same characteristics, such as: of ignorance of facts, inappropriate application of basic probability analyses, and, especially, overconfidence in the model. Forecasts are very often made more inaccurate by overfitting: as noise in big data often increase faster that the signal, pattern are harder to find and it is easier to confuse noise for signal.

Luckily the author set out to write a book about possible solutions to these potential failures. In particular, Silver identifies three solutions that can help any motivated reader to improve one's

predictions. First, use probabilistic predictions and continue to revise the prediction to account for new information; this is also called Bayes' theorem. It works by "*getting closer and closer to the truth as we gather more evidence*". Second, make better prediction by having skin in the game, having something dependent of the outcome of the decision, and using feedback loops to quickly test your hypotheses and adjust the model accordingly. Thirdly, average many predictions. Averaging expert's forecasts yields better predictive results than using the forecast of "an average individual", with the effect being 15% to 20% better.

Silver's book is full of anecdotes and charts to get the author's ideas across and it contains many other ideas and examples. These range from changing our mental model we use while predicting to the advantages of humans over computers in the chaos theory, from the necessity of understanding context to the story of human/machine rivalry materialized in a chapter dedicated to chess, and many others that I encourage you to discover.

Overall, it is a useful book for a first introduction to our biases, and for a reality check regarding our boundaries (see subjectivity and irrationality) and for those interested in answering big questions using predictions. Due the numerous examples contained in this book, my recommendation is to read the book feeling free to skip chapters that seem superfluous to your tastes.