Baysian Machine Learning

In the last blog, we focused on the fact that in order to make optimal decisions under uncertainty, such as investment decisions, you need a future scenario (the return expectation for a stock) and the associated probability of its realization.

The calculation of the probability consists of the sum of two types of uncertainty, data uncertainty and model uncertainty. We have been able to quantify data uncertainty up to today, but not model uncertainty. It involves the calculation of all the possible model specifications capable of explaining a given set of data to find the probability that the selected model is the correct one.

Bayesian machine learning models have enabled us to do just that, but why exactly?

To understand it, we must go a step deeper and look at Bayes Theorem and the concept of maximum likelihood. It can get a bit technical, but we do our best to keep our feet on the ground.

Bayes Theorem

Below is Bayes Theorem, a way to do proper hypothesis testing or inference, when faced with uncertainty. Even though many within the fields of natural and social sciences agree that this is indeed the right way to observe the physical reality in which we live, historically there was a problem or actually two problems with the below theorem.

Likelihood
How probable is the evidence given that our hypothesis is true?

Prior
How probable was our hypothesis before observing the evidence?

$$P(H|e) = \frac{P(e|H)P(H)}{P(e)}$$

Posterior
How probable is our hypothesis given the observed evidence?

Marginal
How probable is the evidence under all possible hypotheses?

The first problem is the prior knowledge (red text). Incorporating prior knowledge required scientists to state their minds, which is not easy, and furthermore to formalize this knowledge. The second problem is the marginal probability (blue text) which requires the estimation of all possible hypotheses or models given the observed data, an exponentially difficult estimation procedure with potentially billions of parameters to estimate.

The solution to the two problems became to naively assume that we have no prior knowledge of the problems we try to solve, removing the prior term in the numerator, and also not treat problems as probabilistic, thereby reducing the marginal probability in the denominator to just a normalizing factor, which can be removed. The only term remaining on the right side of Bayes theorem is the likelihood and as such we can infer things about the world by calculating only the maximum likelihood. But to get there we made a lot of questionable assumptions!

Maximum likelihood

Maximum likelihood basically means a procedure that best-fits the parameters of a model to a given dataset. This is problematic when the data used to fit the model is different from the data the model is supposed to explain. Dynamic data, such as most financial data, may not contain a good representation of the future and since maximum likelihood models only use data, the ability to extrapolate into the future is limited.

The maximum likelihood models do not incorporate any prior knowledge that we might have and at the same time the model is forced to provide a best fit. This combination makes the maximum likelihood models prone to finding correlations and not causality.

The key to solve this problem lies in the concept of Bayesian machine learning, which combine data, domain knowledge and probabilistic modelling. By incorporating prior knowledge of the problem at hand, coding it as a statistical distribution and then combining it with the information in data, significantly reduces the risk of fitting the model to statistical only truths (hypothesis generation and not hypothesis validation).

On top of that, Bayesian machine learning models are allowing the model to doubt its conclusions, forcing it to disregard spurious correlations and quantify the validity of its predictions. The combination of all this results in a model that actually has an understanding of the environment in which it operates.

In order to harvest the benefits of machine learning, we as humans need to be smart in our use of this powerful new tool. Most popular models from machine learning today, like neural nets, are merely fitting a line to existing data i.e., an advanced linear regression.

Regardless of how complex or deep the network is, the network does the exact same thing. You may configure it differently to get representations in different layers and make it more effective, but the crucial part is how the model must relate to the data it is trained on, as well as the variables it tries to explain. Are these static or variable? A simple perception model only gives you one answer that fits your model without questioning it, much like classical maximum likelihood models.

What you can do with Bayesian predictive inference machines using probabilistic programming is to give the algorithm room for doubt and the possibility for the model to disproof your assumption. In this way we minimize the risk of fitting our model to the spurious correlation in data and maximize the probability that we fit our model to the true causal dependencies corresponding to the physical reality that we live in.

The new doctrine

Mathematicians, statisticians, physicists and investment professionals have been indoctrinated with models that were never meant to solve complex real-life problems. Models that relied on strict assumptions in order to fit a reality that never existed, but so far this was justifiable since there was no alternative. Today there is!

Baysian machine learning provides us with the ability to make optimal decisions in a world of uncertainty. Today we can model the world without relying on unrealistic assumptions resulting in statistical-only results.

Yes, investors are forced to step into unknown waters, educate themselves in fields not formerly required by the finance community and accept that computers can now solve certain tasks much better than humans. However, In AI Alpha Lab we believe that those who take the leap of faith, educate themselves and incorporate machine learning, will be greatly rewarded. They will be at the forefront of active investing in the coming decades, taking more informed investment decisions and be armed with the knowledge of what they don't know!

Be sure to read our next blog as we show that an uncertainty model enables you to invest like a loser – and why that's a good thing!

"It ain't what you don't know that gets you into trouble.
It's what you know for sure that just ain't so."

- Mark Twain

Disclaimer

This blog has been prepared by AI Alpha Lab ApS (AI Alpha Lab) and contains information and inspiration for the reader. The blog must not be considered investment advice and cannot be relied upon as a basis for a decision to buy or sell securities (or refrain from doing so). AI Alpha Lab assumes no responsibility for decisions made on the basis of information in the blog.

This blog is provided for information purposes only and does not constitute, and shall not be considered as, an offer, solicitation or invitation to engage in investment operations or as investment advice. If investment strategies or portfolios of securities are mentioned, the reader must be aware that these are not necessarily advantageous for all investors, and that AI Alpha Lab does not know each reader's individual characteristics such as knowledge, investment experience, finances, risk preferences etc.

AI Alpha Lab has taken all reasonable precautions to ensure the correctness and accuracy of the information in the blog, as it is based on information obtained from sources believed to be reliable. However, the correctness and accuracy are not guaranteed, and AI Alpha Lab assumes no responsibility for any errors or omissions, including typographical errors.

Any information or opinions are issued by AI Alpha Lab as of the day of publication and are not updated or changed subsequently, unless otherwise clearly stated. The reader is encouraged to seek individual advice on their own investment matters.