MODULE 6 UNIT 1
Video Set Video 1 Transcript

Module 6 Unit 1 Video Set Video 1 Transcript

STEFAN ZOHREN: Welcome back, everyone. In previous videos, we already spoke a little bit about machine learning applied to finance in the context of limit order books and market microstructure data. Obviously, there we have a lot of data, as we discussed, and there's also some more structure, so that's obviously one of the best areas to apply such more advanced machine learning techniques.

In this video, I wanted to cover some other recent advances on how we can use some of the more state of the art machine learning techniques, such as transformers and attention, in the context of traditional quant finance, trading strategies in particular, and the interplay of momentum and reversion and how one can use such powerful techniques also in a much data-scarcer environment where we have, for example, only daily data.

Transformer models

ZOHREN: So, you might have already heard about transformers or attention, which are some of the techniques which make those large language models such as ChatGPT so powerful. And while they are applied in a completely different context here—we don't apply them to text at all. There's no text. It's just financial data. But might still be useful to understand the concept in terms of languages.

For example, when you look at attention in a language model, you can see how the model assigns certain attention rates between different words, which assign a little bit of the importance different words have on each other. If you have, for example, a sentence like, "I'm sitting on the river bank", then there will be a very strong attention rate between "river" and "bank". And that helps the model to figure out that the word "bank" actually is a "riverbank" and is not "bank" as in the sense of financial institution. So, this is essentially how this language model works, and this attention makes them so powerful that they can understand these slight nuances, that this is a riverbank and not a financial bank.

Applying transformers to financial data

ZOHREN: Similarly, we can apply the same techniques to financial time series. Now I don't have any words – I just have returns going up and down or momentum indicators or any other technical indicators – but now the model is trying to pay attention and to basically make decisions at any moment in time, and it can refer to different points in the past.

And we have this paper, which is called "The Momentum Transformer", which utilises those techniques and tries to effectively tune around between momentum and reversion indicators. And this is a very hard task because, generally speaking, I have trending markets—so, time series momentum where markets which have been going up are more likely to continue to go up. But, obviously, one big issue is when the market is about to turn, and I have potential crashes. And it's important for the model to figure out when those momentum turning points are, when it has to switch from trending to following more of a reversion pattern.

And that's where we have found that such attention-based models are quite powerful. Not only are they very powerful in terms of performance, but you can also get some interpretability out of those models by looking at the attention rates. And we can see, for

IN COLLABORATION WITH getsmarter™

example, when we run through the 2020 COVID crash, that after the crash, the model keeps on attending to the dip in the market. So, it can decide that some long-term trend indicators, which span a big period before the actual crash are probably not relevant anymore because we are in a completely different market regime. And it keeps on attending to this dip in the market as we go along, and thus learns how to focus more on short-term predictors when making its decision. And once we are further away from this crash, it will start to adapt again to some of the momentum indicators.

And this is actually quite powerful because some of the traditional momentum estimators, some we have already got to know in some of the tutorial sessions like MECD filters, have actually performed quite badly during the COVID crash, which was very, very fast. So, some of these estimators, when the market crashed, they updated slowly and effectively traded out at the bottom of the market, and then the market very quickly rebounded. While those transformer-based models were very fast to effectively reanalyse the new situation and then updates its decision based on that.

Conclusion

ZOHREN: So those are very, very powerful techniques which come from some of the more recent advances in machine learning and, in particular, some of the advancements in large language models without actually looking at language at all. You're just applying those techniques to time series model. So, this is a very interesting area, and it's definitely a space to follow and to maybe do some further reading in.

So, thank you.

SPEAKER: Did you understand all of the concepts in this video? If you would like to review any of the questions, click on the corresponding button.

IN COLLABORATION WITH getsmarter™