

Hi, I tried to compare Numerbay models.

Why? Unlike buying art, from Numerbay models we usually want the same thing: to perform well. The Numerbay marketplace is not very friendly in this regard, as it basically shows hundreds of models, which can lead to a choice paralysis. But all we want is the one best model, right?

So I built a script that fetched some data for the best models on Numerbay (based on their 1 year average score).

My metrics are:

annualized score ($2mmc+0.5$

corr, full year average, i. e. not filled with zeros)

score for 1 NMR (how much score is expected for every 1 NMR paid for the model's predictions)

trust score (how the model owner trusts the model, as shown by their stake in this model / total stake; some models are pretty good, but their owners have tens of Numerbay models, each one with small stakes, probably hoping for at least one model's luck, I would not select one model randomly performing well)

days in last year (how time-tested the model is, maximum is 365 days)

required stake (how much you need to stake to break even; considers annualized score and price)

The models are sorted by their dragon score:

(\ln of "days in last year" * "annualized score" * "trust score") / price

Of course these are not the only right metrics, there are so many things to look and not to look for. Feel free to develop your own metrics and scores! For example one that weighs more recent rounds would be interesting.

So here it is:

[

Snímek obrazovky 2024-05-03 164959

847×641 24.1 KB

](<https://forum.numer.ai/uploads/default/original/2X/3/3fd0f0b59fb6fd1607dc5fa2dc27516e137c7fac.png>)

I highlighted 6 models, which I consider to be the most interesting, worth watching and maybe subscribing to (disclaimer: 50NN_300 is mine).

What I learned from doing this:

- Many prediction buyers are probably losing money, because the required stakes are pretty high. Especially those of models whose owners don't stake themselves.
- Many models, including some popular ones, with very low scores.
- A common strategy of listing many models with almost no stakes. As I already said, I do not consider this an honest strategy.

Thanks for reading this post and maybe thinking about Numerbay model evaluation for a while. I think it's an interesting topic and feel free to discuss my ideas or post your own.