1. Dia (nevem meg a sztaki)

---

2. Dia (Cím)

I am going to talk about distributed evolutionary selection, and about how we tried to find the weights of a given neural network using this method.

So it is a distributed neuro evolution – which is quite an old technic, but we did a couple of interesting things and I try to show that this can still be a useful approach in some cases.

3. Dia (Overview)

To demonstrate the usability of this approach we chose a time series which comes from the forex market – this dataset is the Euro USD exchange rate from 2 thousands till now in a minute resolution – which means that we had 5.6 million observations.

4. Dia (EUR/USD Exchange rate)

It is an ordered time series data.  
The future of the exchange rate is unknown, and unpredictable.  
**~~We can not control the price.~~**  
~~In other words, we are too small players in a huge market.~~  
It is hard even if not impossible to execute experiments or create controlled tests and hypotheses because the circumstances are always changing.

5. Dia (Sötét hátterű dia)

So what was the question? Can we follow the movement of the price? Can we take advantage of this? And if the answer is yes then how does look a system like this?

So I wanted to write an algorithm that is able to find a strategy (which is profitable), with the help of a Machine Learning algorithm.

Isten ments, hogy igereteket tegyel!!!! Nem tudjuk mit hoz a jovo, epp ugy mint az analizalt tozsdei grafikonjaidon! ☺

6. Dia (Motivation and propositions)

Instead of trying to predict the future of a given time series at any given point in time – like the animation shows – we chose another approach.

7. Dia (Motivation and propostions)

We also could have defined this as a classification problem, if we annotate the data where or when should we buy and where should we sell, then we have to train the algorithm to find some relation between the labeled data and the input, where the input is the past of the time series itself.

But there are a lot of problems with this approach.  
We drop a lot of useful information if we do not use all of the data.  
Each point should belong to one of the three categories (buy, neutral, sell)  
It doesn’t matter how we want to label the data it always will be problematic.

8. Dia (Our approach)

Instead of annotating the low entry point of trades and high exit points of trade manually, we wanted to find these points automatically and algorithmically.

So our objective function was to maximize the distance between the buy prices and sell prices, based on the price history itself.

I would like to give you an example.

For example, in this case, this solution would have very good results.

The green area shows the profit of the given trade.

This is what we would like to maximize.

In this case, there wouldn’t be any better. (or …)

9. Dia (Our approach)

But I have to highlight the fact that this is also a good solution even if it is not as good as the previous one.

10. Dia (Our approach)

Many other – and even better solutions exist – and this is what we are trying to find programmatically.

11. Dia (How did we define the task?)

Itt az a szöveg ami a képernyőn van.

12. Dia (Amin a Neurális Háló van nagyban ábrázolva)

I cite this picture only to get a better insight how does the neural network look like.

The thickness of the lines represents the weight of the connection, the thicker the bigger, while the color of the lines is the sign of the weight (blue is negative).

What we are trying to find is the best weight of the Neural Net.

The bottom lines (blobs) are the input of the neural network. In this case, the lookback window was 30 (thirty) so we have 30 inputs, but this number can choose a bigger lookback window. Of course, if we have more inputs or we choose a longer lookback window then the neural network will be more complex.

While we do not know what are the best weights, we try out a lot of possible candidates and choose the best one from time to time.

13. Dia (Amin Animálva végig tolom a hálót az idősoron)

So I send the candidate neural net to the worker – which processes the given time series using the sent neural network – just like a convolution (or a function)

The input of the Neural Net is the last 30 values of the price at any given time.

The output of the Neural net is a Real Number.

It is important to understand that I do not want to estimate or approximate the original time series with the neural network at all.!

Instead of that, I want to create a function that „finds the inflection points of the original time series”. And this whole thing comes from the past of the time series itself.

14. Dia (Amin már megvan Sárgával az NN kimenete)

When this number is bigger than 0 then we open a buy position and when it is less than 0 then we Close the opened Buy Position and sell the given instrument.

If the output of the Neural Net is greater than 0 then buy at the given price (let this mean buy)

If the output of the NN is less than 0 then sell at given price.

15. Dia (Amin azt mutatom hogy veszteség is lehet a kereskedésben)

If the sell price is less than the buy price then we call it a loss of the given trade, as I highlighted with the orange color in this picture.

16. Dia (Amin azt mutatom, hogy több kereskedés is van az ábrán)

Of course, there are a lot of other trades in this figure, because the setup trading rules are mutually exclusive.

17. Dia (Amin 1000-ből 5.6 millióra nagyitok)

As you can see we have 1000 observations, but altogether we have 5.6 million observations.

And what you see on the bottom is still a fifth of the entire dataset.

The whole dataset is shown on the top plot. ~~So I hope that you have got the point.~~

18. Dia (How did we find the proper weights of NN?)

We search the weight of the neural net with evolutionary search method.

And our motivation was to decrease the computation time.

19. Dia (Open Stack screencast)

Skippelve

20. Dia (az első animált Arhitektúra ábra)

So we take the cloud

Create a lot of Virtual Machines (as many as the population size)

We also have a Driver, which initializes the whole cluster (so we do not take care of this)

It creates a Neural Network with a given architecture.

Not just one but many, and of course with different weights. (Well of course these are just random guests)

21. Dia (a második animált Archtektúra ábra – neurális hálók szétosztása)

1. The driver takes care of the selection, the crossover, and the mutation
2. And the workers do the computation-intensive part of the task

22. Dia (Architektúra ábra amin a Crossover és a Mutáció van)

1. What do the workers do? They compute the output of the f(x,w) function for every observation and compute the fitness score of the given solution – which is in this case the total gain of the trades. (let’s say the profit)
2. It takes a while for every worker to get the results. For example to get the result of every F(x,w) and based on that value to calculate the fitness score of the given solution.
3. But immediately after a Worker finishes its job, it sends back the fitness score to the driver, which holds the given solution in its memory.
4. Sorts the solutions by their fitness scores.
5. We apply some sort of selection mechanism. (For example, select the best or the best two)
6. Apply some crossover on this (or not)
7. Mutate the new candidates a little bit
8. Distribute the new candidates again among the workers
9. Until it reaches some predefined stop criteria.

23. Dia (Repeat)

Repeat this process until it reaches some predefined value (while hopefully we find better and better solutions)

24. Dia (Train test split)

* Of course overfitting can occur so we tested this method in two ways.
* First we split the data into a train and test sample.
* Second we split the data into a train, a test, and a holdout sample, and we continued the evolutionary search as the new data was coming in. (arrived)

25. Dia (How does the equity change through evolution?)

Let me show you an animation. The orange line is the equity (or the gain of the trades). This is what we want to maximize. While the blue line is the given price of the given instrument.

This short animation shows us how to change the Equity (orange line) through evolution.

The plot shows only a short slice of the whole time series (by the way)

But you can see that the algorithm could find a very good solution.

26. Dia (Results 3/1)

Let me show you an animation. The orange line is the equity (or the gain of the trades). This is what we want to maximize. While the blue line is the price of the given instrument. Now the interesting thing is that the learned relationship between the past and the trading action remains profitable on the unseen test sample, without any modification of the neural network. (The vertical line shows the end of the training sample)

We did a lot of tests. We examined the distribution of the trades, we compared the train and the test distribution.

We want to show that this sort of Neuro evolution can work on financial dataset like EURUSD with good results.

We were able to make the computation faster using cloud infrastructures.

27. Dia (Results 3/2)

* as we increase the population size which is equal to the number of Virtual Machines decreases the computation time.
* For example, if we set the population size to 40 then the speed up was around 38 times faster (but this value strongly depends on the size of the data)
* but the speed up does not mean that the bigger population size leads to a better solution (better fitness score)

28. Dia (Result 3/3)

Probably it is not so obvious how sensible or stable are some solutions for the variation of the weights.

So this second picture seems to be a little fuzzy but you can imagine the distribution of the fitness scores.

For example, if we watch the orange or the blue dots then we can see that the candidate solution fluctuates around some values.

This horizontal section developed because we always keep the best solution of the previous generation in the new generation as well.

29. Dia(Result 3/3)

But why was so important for us to increase the processing speed of the computation via parallelizing this searching method.

Because it is a metaheuristic search and we had a couple of hyperparameters that influence the result.

I can mention the architecture of the neural net, (for example the number of neurons) or the rate of the mutation of the weight.

So we were able to try a lot of setups to find a better solution.

And we also were able to run a longer experiment.

30. Dia(Future works, plans)

My plan is to replace the Fully Connected Multi-Layer Perceptron with a Multi Layer Convolutional Network (probably I can reach an even better result with this) (while the way of searching the weight of the net remains the same)

Ha olyan kerdest kapsz, amire nem tudod a valaszt vagy nem tudsz jo valaszt, esetleg ratapint a megoldasod gyengesegere valaszold a kovetkezot:

I think this is a good idea. I think it is worth considering. Thank you for raising this topic.