

Project proposal

1. Project name

Analysis of the financial time-series based on its decomposition

2. Team. Team members. How tasks will be allocated among them?

Karpov Gleb

Ponomarev Evgeny

Vetsler Denis

Zankin Vitaly

3. Background. Why is the problem worth attention? And how linear algebra is going to solve it?

We consider some common patterns and methods to extract the associated components from a time series. It probably will help us to understand the time series better and also we hope that it can improve volatility class forecasts. Why? Because the input “signal” is too noisy and decomposition of input values for classifier is going to decrease such noise level in order to solve the problem.

4. Problem formulation Define the problem formally, specify necessary equations/models.

Improving the precision value of a prediction (1 minute future) volatility class via changing the input vectors for classifiers. How we will change is that exactly where NLA will help us. We will use decomposition of time series for the feature values minutely. It means that the classifier will get decomposed values (only trend for example) on his input.

5. Data. Tell us about the datasets you will be working with here.

We will use the given order log from Stock server. That order log is already prepare to work with it. Features are already selected. We have training set of the feature values with corresponding values of volatility. Also, we know the precision of volatility class forecasting based on our initial feature values input.

6. Related work. What has been done on this topic? Google Scholar might be helpful :)

7. Scope. What will be the end result of the project? Describe what phases of work will be undertaken.

In the end of work we will get the evaluation of how the different decompositions of the input financial time-series effect on xgboost classifier precision on our data.

steps:

1. survey of what decompositions are exist and which of them can be interesting for our case
2. choosing several (2+) decompositions for implementing
3. implementation and verifying decompositions
4. integration in our classifier
5. tuning parameters of classifier and decomposition (if any)
6. getting result

8. Evaluation. How do you measure the performance of your solution? It is also a good place for including baselines.

Now our team is working on prediction of volatility of shares by using history of sell/buy orders.

The idea that there are several different trends on the market and using decomposition of input time-series will give us an improvement in quality of prediction.

So, our final goal is to improve prediction quality (precision and recall) via decomposition of input data.

Current quality of the prediction for using three equally like classes (high, med, low) of volatility is 67% (average precision)

References