The aim of the project is to develop a novel method of commodities prices forecasting. The proposed method is a mixture of so called "symbolic regression" and methods known from the Bayesian approach.

First of all, modelling of commodities prices is a hard task. However, it is crucial for markets, economy and finance. Usually, it occurs that it is hard to construct the model, which would produce more accurate forecast than, for example, the naïve method. Therefore, searching for new models is an important research task. Simultaneously, it is a necessary and useful task, and it can lead to the better understanding of which processes take place in modern economies.

The next problem is that usually there exist a lot of factors potentially influencing the development of commodities prices. This leads to the necessity of variable selection, if a model is fixed. This is not a trivial task, and by itself is an important obstacle.

The proposed method is based on so called "symbolic regression". This method resembles a bit the Darwinian theory of evolution. Indeed, first one has to initialize some set of models. Of course, these models are denoted as some functions, i.e., through mathematical symbols. In the next step, similarly like in the theory of evolution, two processes can happen.

The first process is "mutation". In other words, the functional form of a given model can slightly change. The second process is "crossover". In other words, a part of the function describing the given model can be taken, for example, the first part of its symbolic (mathematical) notation. Next, the ending taken from the mathematical notation of another model can be appended to it. This way, a "child" is obtained, which is some new symbolic notation, representing a new model.

It is expected that by repeating these steps the "best" model will "evolve".

The Bayesian approach can be described in the following way. Assume that we estimate some unknown parameter, which has some prior distribution. When the set of observations expands (for example, time passes and new information comes from the market) this distribution can be re-estimated – and the posterior distribution is obtained. In other words, starting from some rough estimations on the interesting parameter, we can expect that, as the new additional information comes, we will be able to "shrinkage" our estimations.

Why then not to assume that some "parameters" describing "mutation" and "crossover" in "symbolic regression" can also be described with the Bayesian approach – not like it is usually done: to assume that they are fixed and not-changing? It seems that no attempt to join these two methodologies has been taken yet.

In this project an attempt to join these two approaches will be taken. Also, the suitable algorithm will be constructed and implemented. Next, the developed method will be tested for forecasting various commodities prices. From the economical point of view, it is interesting to look on how the developed algorithm selects particular variables. In other words, which role it ascribes to various, potentially important factors influencing the development of commodities prices.