



WHAT MAKES OUR APPROACH TO AI AND ML DIFFERENT?

Advanced utilization of artificial intelligence and machine learning at Quantumrock

WHAT MAKES OUR APPROACH TO AI AND ML DIFFERENT?

Advanced utilization of artificial intelligence and machine learning at Quantumrock

This document summarizes the main components of the AI/ML approach developed and actively used at Quantumrock. The main emphasis is placed on distinctive elements of our approach - our know-hows and differentiators. To be able to make clear statements we did not try to avoid technical details.

Why Machine Learning?

We consider the use of ML for the generation of trading strategies as complementary to the construction of strategies by domain knowledge experts. Both approaches have their unique strengths. Domain knowledge experts can generalize their knowledge about one domain and transfer it to another one. ML methods can perform an exhaustive search over a huge number of potential patterns in the data and, in this way, to find nonobvious, nonlinear, multidimensional and fuzzy patterns that are hard to grasp by the human mind. We benefit from both approaches not just by using them in parallel but by combining them into one synergetic approach.

Going beyond “vanilla” Machine Learning

Customized Objective Function

One of the main tasks of ML is supervised learning. Within this task, ML aims to find (learn) a function that maps from a given (usually previously unseen) input into a corresponding output. The function is derived from a data set that can be imagined as a set of examples of the wanted mapping (a set of input-output pairs). Inputs are usually given as vectors consisting of numerical components and targets are either classes (in case of a classification problem) or numbers (in case of a regression problem).

One major class of the problems that we solve at Quantumrock has a very similar structure. We have data sets that consist of pairs of features (financial indicators) and targets (price changes of financial instruments). However, predicting price changes (solving a regression problem) does not serve our purpose, because knowing the expected price change for a future period alone does not allow us to decide an optimal position (allocation)

for this period. For example, we might have periods with the same expected price changes but very different volatilities (dispersions) which means that these periods require different positions (allocations).

A proper solution to the problem, that we have developed at Quantumrock, is a direct calculation of the position for the future period based on the features associated with the prehistory of the period. Within this approach, the values (positions) generated by the searched mapping should not be close or similar to the targets (price changes), as in the standard regression problem, and, therefore, the standard measures of model quality (mean squared error, mean absolute deviation and so on) cannot be used. Because of the same reason the values generated by the model cannot be interpreted as "predictions" anymore. The proper redefinition of the objective function is an elegant solution that allows us to use all the well-established ML methods (Linear Regression, Neural Networks, Decision Trees and so on) and, at the same time, to directly optimize what we need from the financial point of view (i.e. positions that maximize the Sharpe ratio).

Time Series Modeling

Another essential difference between the problems that we address at Quantumrock and problems considered "vanilla" ML is that we use (multivariate) time series as input, instead of numerical vectors of a fixed length. A naïve compression of a time series into a vector of predetermined features (i.e.

technical indicators, momentum, volatility or other statistical properties like skewness or kurtosis) allows us to transform the strategy development problem into a standard regression problem and to apply the full power of ML. However, this approach has a severe drawback. It might be the case that by using manually defined features we lose essential information about patterns present in the input time series. This means we might end up in a situation where we apply advanced and powerful ML methods to a poor input and, as a consequence, get poor results. To resolve this drawback or, in other words, to remove the bottleneck of bad features, we use the above-mentioned ability of ML to find complex, non-trivial patterns in order to construct new features. In more detail, we, at Quantumrock, have developed a set of methods that allows us to determine from a raw time series (a sequence of price changes) how this time series should be compressed into a vector of features such that the most relevant information for our purposes is not lost. The main idea behind this approach is that the historic price changes (raw input) are extended by their temporal positions which are explicitly used by the model, so that the model does not need to learn it from scratch.

Heavy use of Mathematical Statistics

In most ML problems it is easy to construct a simple reference model that has some low but still statistically significant predictive power. Such a simple initial model is already somewhat useful. Any further modeling work is dedicated to the improvement of the

model's accuracy (an increase of the predictive power) and, therefore, the usefulness of the model. The situation is different in the case of ML modeling for algorithmic trading, mostly because of a high noise-to-signal ratio in the price data (markets are indeed very efficient) as well as unusual objective functions. Firstly, even a positive performance of a model (in terms of average profit, Sharpe ratio and so on) might be not very statistically significant, meaning that positive results might be easily caused by luck (random, unsystematic fluctuations in the data). Secondly, even a strategy with positive and statistically significant results might be useless or even harmful (it might just lose money less efficiently than a random strategy).

To address the above-described problems, we, at Quantumrock, apply a broad range of non-parametric tests of statistical significance at different stages of the modeling process (starting with the feature selection and ending with the final performance of the strategy).

Strong Focus on Overfit Prevention

Because of the above-mentioned high noise-to-signal ratio in financial data we might easily end up with an overfitted model – a model which by mistake interprets random noise as real patterns. To prevent overfitting, we have developed a set of protection mechanisms that are based on the following components: (1) we search in the domain of models with a relatively small number of parameters, (2) the model structure is carefully selected based on domain knowledge to provide a

proper inductive bias, so that we search only in the space of realistic patterns, (3) we rely on regularization techniques to flexibly suppress excessive model complexity based on the amount of data and its noise-to-signal ratio, (4) we conduct model training, testing and validation with a proper split of the data (the final version of the model is always tested on the data set that was not used during the model development process), (5) the found models are mixed properly such that the remaining noise is canceled out further (at the stage of assignment of weights to the strategies we also apply regularization techniques preventing overfitting).

Benefitting from the Deep Learning revolution

The recent progress in computational power and the amount of data stimulated huge advances in theory and practice of training extremely large models (Deep Learning). Usually, Deep Learning Models are operating in the domain of images, videos, audios and larger text corpora. In spite of the fact that Deep Learning deals with large models and we, as has been mentioned earlier, focus on small models, it is still possible for Quantumrock to benefit from the Deep Learning revolution. In particular, we can extract components that are not per se bound to the number of model parameters. In Quantumrock's modeling process, we benefit from the following components of Deep Learning: (1) methods for proper treatment of structured data (sequences, grid data, sets, graphs), (2) efficient model optimization methods based on

backpropagation, (3) efficient libraries for operations on multidimensional arrays (tensors).

Automation

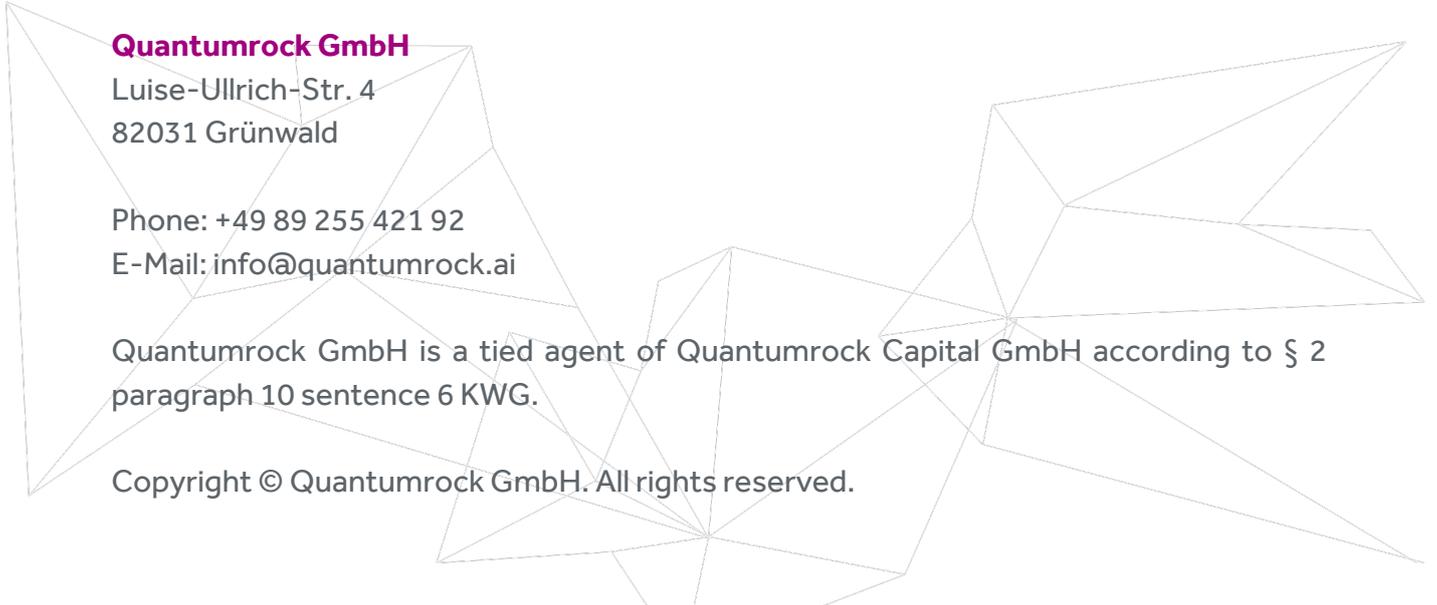
One of the ways to achieve a stable positive performance is to use a larger number of weakly correlated and properly weighted trading strategies. To practically implement this approach, we need an efficient way to generate new strategies that have an added value in the context of already existing strategies. For that reason, we are constantly working on the optimization and automation of the strategy development process. As the outcome of this effort, we have developed an AI/ML system that, after being fed with data, generates trading strategies by automatically performing all the steps of the model development process (features construction, model training, testing, validation and statistical testing).

Summary

In Quantumrock's development of AI/ML system, we distribute our effort over the entire chain of the modeling process to avoid bottlenecks: (1) using domain knowledge experts as well as ML techniques to construct novel, powerful features, (2) construction of ad hoc model architectures to provide a proper inductive bias and to avoid overfitting, (3) adjustment of objective functions to be able to use the full power of ML methodology and, at the same time, to optimize regarding what we need from a financial perspective, (4) extensive use

of mathematical statistics to prove the statistical significance of the results, (5) optimization and automation of the model development and validation process.

This document is for information purposes only and nothing contained in this document should constitute a solicitation, or offer, or recommendation, to buy or sell any investment instruments, to effect any transactions, or to conclude any legal act of any kind whatsoever. Although Quantumrock GmbH ("Quantumrock") believes that the information provided in this document is based on reliable sources, it cannot assume responsibility for the quality, correctness, timeliness or completeness of the information contained in this document. Except as permitted under applicable copyright laws, none of this information may be reproduced, adapted, uploaded to a third party, linked to, framed, performed in public, distributed or transmitted in any form by any process without the specific written consent of Quantumrock. To the maximum extent permitted by law, Quantumrock will not be liable in any way for any loss or damage suffered by you through use or access to this information, or Quantumrock's failure to provide this information. Our liability for negligence, breach of contract or contravention of any law as a result of our failure to provide this information or any part of it, or for any problems with this information, which cannot be lawfully excluded, is limited, at our option and to the maximum extent permitted by law, to resupplying this information or any part of it to you or to pay for the resupply of this information or any part of it to you. Neither this document nor any copy of it may be distributed in any jurisdiction where its distribution may be restricted by law. Persons who receive this document should make themselves aware of and adhere to any such restrictions. In particular, this document must not be distributed or handed over to U.S. persons and must not be distributed in the USA.



Quantumrock GmbH

Luise-Ullrich-Str. 4
82031 Grünwald

Phone: +49 89 255 421 92

E-Mail: info@quantumrock.ai

Quantumrock GmbH is a tied agent of Quantumrock Capital GmbH according to § 2 paragraph 10 sentence 6 KWG.

Copyright © Quantumrock GmbH. All rights reserved.