

# Оптимизация портфеля цифровых активов используя методы машинного обучения

## Данные

Предоставлены следующие данные:

- **symbol** - инструмент (котируется к USDT)
- **time** - время закрытия свеч (UTC)
- **open, high, low, close** - цены
- **market\_cap** - рыночная капитализация актива
- **n\_trades** - количество сделок за последний час
- **volume** - объем торгов за последний час
- **taker\_buy\_volume** - объем агрессивных покупок актива за последний час

Данные поделены на 2 выборки. На первых порах будут данные с 2018-01-01 до 2021-06-01 (тренировочные). В конце проекта будут данные полные данные для сверки итогового результата. Бенчмарком по перформансу принять модель HODL BTC.

## В течение первой-второй недели:

Необходимо преобразовать рыночные данные в фичи и рассмотреть их влияние на будущую доходность (3, 12, 24 часа).

В преобразовании можно использовать моментумы, стандартные отклонения, квантильные отсечения, ранк капитализации и т.д. При применении функций, связанных с квантилями, необходимо соблюдать скользящее окно (без заглядывания в будущее).

Также необходимо создать агрегированный индиктор (взвешенные по кэпу или средневзвешенный) и прикрутить к нему фичи. Далее добавить эти данные к каждому инструменту.

Во время анализа фичей необходимо для себя будет ответить на следующие вопросы:

1. В моменты сильного роста рынка какие активы лучше всего растут? Как ведут себя эти активы в моменты сильного падения рынка?
2. Какой будущий период лучше всего может быть объяснен с помощью данных? Почему?
3. Необходимо вычленив лучшие индикаторы и дать им личную оценку с точки зрения предиктивности. Как вы можете оценить какой период индикаторов лучше всего применим для оценки будущей доходности?
4. Можем ли мы, учитывая значения момента, сказать об устойчивой зависимости будущей доходности и ранка капитализации актива?
5. Можно ли по агрегированному индексу дать оценку того, как сильно упадет рынок в ближайшее время? А вырастет?
6. Имеет ли значение величина стоимости актива относительно остальных?
7. Как ведут себя outliers по моменту/волатильности и их изменениям.
8. Как коррелируют активы между собой? Можем ли мы определить текущую фазу рынка, используя матрицу корреляции?

Следующим этапом можно будет приступить к построению модели машинного обучения и анализом ее результатов.