

Факультет компьютерных наук



Алгоритмический трейдинг: высокочастотные, низкочастотные и глобал макро стратегии

Выполнил студент:

Мельников Артём Максимович

Руководитель ВКР:

Злотник Андрей Александрович, Директор института количественных финансов НИУ ВШЭ

Высокочастотный трейдинг (HFT)

1

Для управления волатильностью криптовалютного рынка необходимы передовые алгоритмы.

2

Своевременное принятие решений зависит от быстрой и точной обработки финансовых данных.

3

Недостаток публичных исследований в этой области подчеркивает важность дальнейшего изучения этой области применения ML.

4

Интеграция машинного обучения имеет решающее значение для повышения точности и адаптивности.

1

Генерация признаков: Разработка и интеграция функций, которые точно отражают текущее состояние рынка, обеспечивая эффективное отражение в нашей модели текущей динамики рынка.

2

Разработка ML моделей: Создание и совершенствование различных моделей машинного обучения для прогнозирования трендов на криптовалютном рынке.

3

Сравнение моделей: Проведение комплексных сравнений для оценки эффективности различных моделей в контексте HFT на срезе актуальных данных.

Общий обзор предлагаемой модели

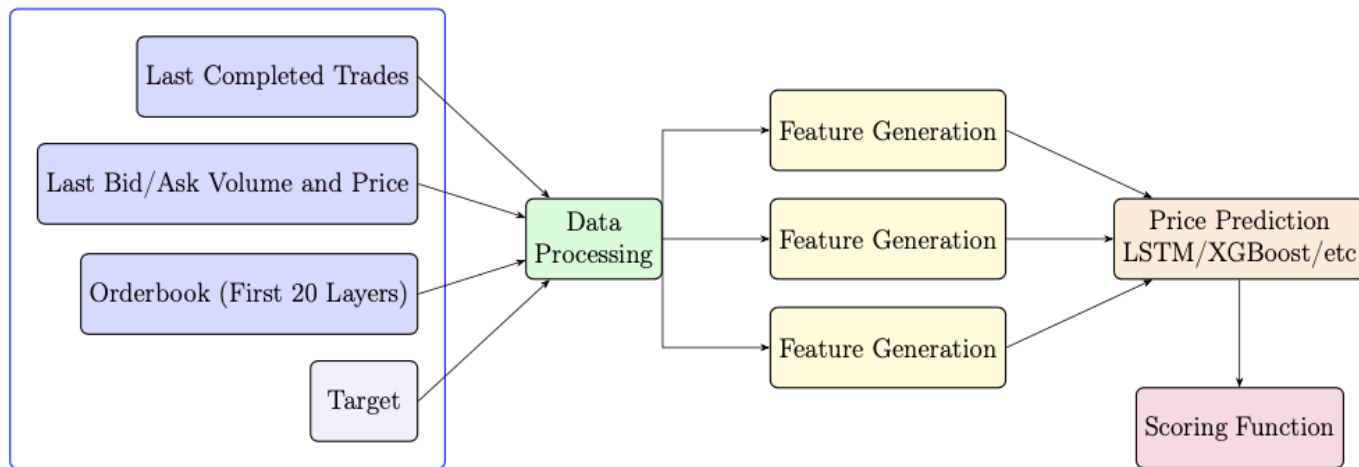
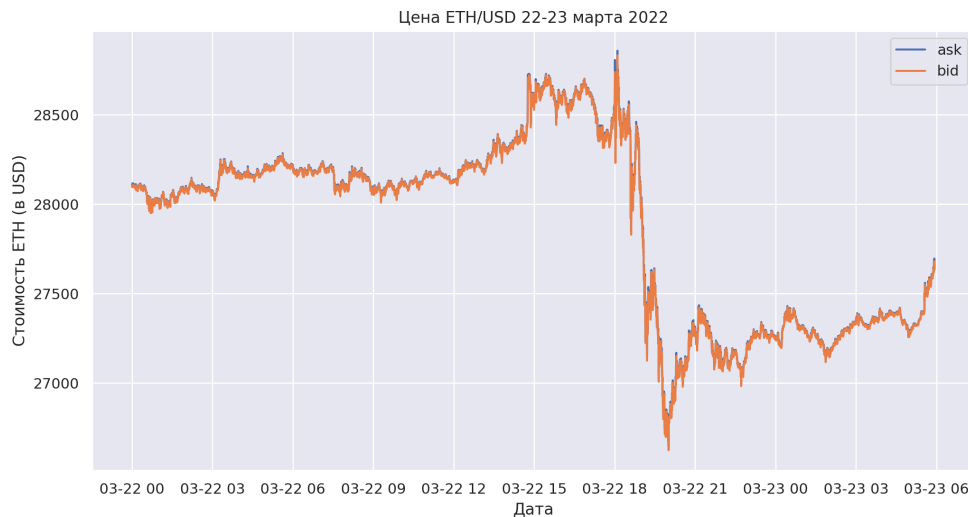


Рис. 5.1: Архитектура высокочастотного ML бота для прогнозирования цен

Данные за март 2023

- Более 10млн сделок в день
- Полная информация о первых 24 уровнях order book
- Постоянно работающий робот-тикер, который получает информацию каждые 3 миллисекунды



- [illegible]

Higher School of Economics
Faculty of Computer science

Гиперпараметр	Предлагаемый диапазон
Количество слоев	1-4
Количество нейронов в слое	50-50
Скорость обучения	0.1, 0.01, 0.00
Размер партии	16, 32, 64, 128, 25
Коэффициент отсева	0 - 0.5
Функция активации	tanh, relu, sigmoid
Эпохи	Основано на ранней остановке
Оптимизатор	SGD, Adam, RMSprop

LSTM

LSTM отличается своей способностью улавливать долгосрочные зависимости и сложные закономерности в данных временных рядов, что необходимо для высокоточного прогнозирования движения рынка.

XGBoost

Благодаря способности эффективно обрабатывать большие массивы данных и высокой точности прогнозирования XGBoost отлично подходит для выявления тонких закономерностей и тенденций в сложных финансовых данных.

Гиперпараметр	Предлагаемый диапазон
Скорость обучения (learning rate)	0.01, 0.05, 0.1, 0.2
Количество деревьев ($n_{estimators}$)	100, 200, 300, 500
Глубина дерева (max_{depth})	3, 6, 9, 12
Min вес в узле (min_child_weight)	1, 3, 5, 7
Коэффициент подвыборки (subsample)	0.6, 0.7, 0.8, 0.9
Регуляризация L1 (alpha)	0, 0.001, 0.005, 0.01
Регуляризация L2 (lambda)	1, 1.5, 2, 3

Для нашего исследования в области HFT мы использовали собственную метрику, разработанную по рекомендации профессионала из индустрии, что позволило нам более точно соответствовать отраслевым стандартам.

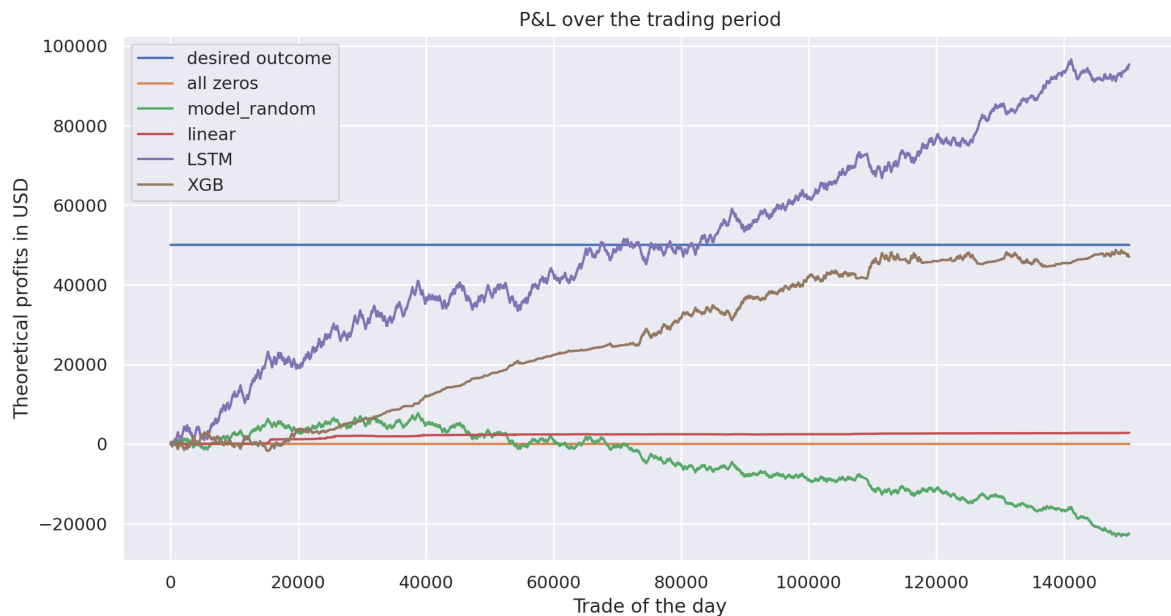
$$\text{metric}(y_{\text{true}}, y_{\text{pred}}) = \frac{\sum (\text{confusion matrix} \odot \text{weight matrix})}{\sqrt{\text{total action count}}}$$

В нашем исследовании линейная модель была выбрана в качестве эталона для сравнения эффективности сложных моделей, таких как LSTM и XGB.

Результаты показали, что, несмотря на более высокую точность LSTM, более быстрая обработка XGB делает ее более подходящей для практических торговых приложений.

Model	Size	Metric	Inference Time (ms)
all zeros	0.070KB	0.0	14.01
ideal model	0.070KB	2825.38	24.89
model_random	0.071KB	-75.91	63.25
linear	2.590KB	9.41	608.06
LSTM	5.000MB	320.36	33969.46
XGB	1.370MB	158.04	5445.39

Таблица 6.3: Final Metrics, Model Sizes, and Inference Times



Обучали на 6 днях, потом график по последнему дню торгов

Не супер честно тк веса confusion matrix могут отличаться от реальности, но хороший PoC

1

Разработка моделей и метрик: Успешная разработка и тонкая настройка моделей LSTM и XGB, а также пользовательской метрики в сотрудничестве с экспертами отрасли для улучшения высокочастотных прогнозов тенденций на криптовалютном рынке.

2

Анализ рынка: Более глубокое понимание волатильности криптовалют, подчеркивающее необходимость адаптивных и надежных моделей. Это исследование также выявило потенциал для анализа рынка в режиме реального времени.

3

Основа будущих исследований: Заложили основу для дальнейшего изучения HFT, подчеркнув важность постоянного исследования и разработки моделей, особенно в контексте обработки данных в реальном времени и более широких финансовых приложений.

Спасибо за внимание!

