

Предсказание цены криптовалюты с помощью временных рядов

Работу выполнили:

Югансон Наталья, Макарова Ксения
НИС БЭАД221

11 ноября 2023 г.

Содержание

1	Введение	3
2	Задача	3
3	Описание и подготовка данных	3
4	Исследуемые методы	3
4.1	Статистические методы	3
4.2	Методы машинного обучения	3
5	Результаты	3
6	Источники	4
6.1	links on literature:	4
6.2	online courses used:	4

1 Введение

Прогнозирование цены криптовалют представляет значительный интерес для многих людей. Это можно делать с помощью временных рядов, что мы и покажем в этой работе. Проанализируем данные за определенный период времени, затем по этим данным построим предсказание и узнаем, насколько наше предположение оказалось верным.

2 Задача

Предсказать цену на криптовалюту используя разные методы и узнать, какой больше всего подходит для нашего случая. Из датасета известно, в какой момент времени была какая цена. После обработки этих данных мы прогнозируем цену на будущее. Следующим шагом мы считаем среднюю абсолютную ошибку (mae) и узнаем корректность нашего результата. Сравнивая ошибки, понимаем, какой метод привел к лучшему результату.

3 Описание и подготовка данных

Датасет содержит данные о цене криптовалюты в течение 547 дней и насколько цена менялась за этот день. Все данные разделяем на тренировочную и тестовую часть. На тренировочной мы будем считать, что должно получиться, а на тестовой проверять, насколько хорошо мы предсказали. Также нужно сформировать данные о предыдущем значении ряда (лаговые признаки) и среднее значение ряда по дню (агрегированные признаки).

4 Исследуемые методы

4.1 Статистические методы

Все статистические методы подходят для стационарных рядов. Поэтому первый шаг это привести ряд к стационарному, затем с помощью линейной регрессии строим прогноз для стационарного ряда. И применяем обратные преобразования к прогнозу. Для нашего датасета mae получается приблизительно 800. Попробуем еще!

4.2 Методы машинного обучения

Существует много методов машинного обучения для прогнозирования, мы использовали бустинг. Для прогноза в частности использовали библиотеку catboost, модель CatBoostRegressor. Mae получилось 227. Первый график - это спрогнозированная цена на последние 7 дней, а на втором показано, насколько отличается наше предсказанная цена от той, которая будет на самом деле.

5 Результаты

Нет универсального метода, который всегда считал бы цену криптовалюты с минимальной ошибкой, но для нашего случая метод с градиентным бустингом, а именно его частным случаем - моделью catboost оказался лучшим. Этот метод мощен тем, что в процессе работы построил

композицию из слабых моделей, деревьев решений. Он способен выявлять сложные нелинейные зависимости в данных, что невозможно сделать вручную. Думаем, текущий результат можно было бы улучшить, если учитывать контекст данной задачи - политический (войны, конфликты), социальные волнения, период анализа, выбранный горизонт прогнозирования и в целом особенности данных. Конечно, на будущее стоило попробовать комбинацию нескольких методов или тщательнее настроить параметры моделей.

6 Источники

6.1 links on literature:

1) Seçkin Karasu, Aytaç Altan. *Prediction of Bitcoin prices with machine learning methods using time series data*. IEEE, 2018. (дата обращения: 08.11.2023)

2) Murali Mallikarjuna Rao Perumalla, Sanjay Kumar Singh, Aditya Khamparian. *Cryptocurrency price prediction using traditional statistical and machine-learning techniques*. The Journal of Finance, 2021. (дата обращения: 09.11.2023)

3) Ahmed Ibrahim, Rasha Kashef. *Predicting market movement direction for bitcoin: A comparison of time series modeling methods*. The Journal of Computers & Electrical Engineering, 2021. (дата обращения: 09.11.2023)

4) Betul AYGUN. *Comparison of Statistical and Machine Learning Algorithms for Forecasting Daily Bitcoin Returns*. The European Journal of Science and Technology, 2021. (дата обращения: 10.11.2023)

6.2 online courses used:

some lectures from fcs course (3d year)