

ПРИМЕНЕНИЕ МЕТОДА МНОГОМЕРНОГО МНОГОЭТАПНОГО ПРОГНОЗИРОВАНИЯ НА ОСНОВЕ МОДЕЛИ ГЛУБОКОГО ОБУЧЕНИЯ LSTM ДЛЯ ВРЕМЕННЫХ РЯДОВ ЦЕНЫ НА БИТКОИН

Дандамаев С.А.

*Национальный Исследовательский Университет Высшая Школа Экономики,
Москва, Россия
ssdandamaev@yandex.ru*

Сизых Д.С., Сизых Н.В.

*Национальный Исследовательский Университет Высшая Школа Экономики,
Москва, Россия
D.Sizykh@gmail.com, sizykh_n@mail.ru*

Аннотация. На данный момент методы, основанные на архитектуре глубокого обучения LSTM, показали лучшие результаты в прогнозировании цен на криптовалюту. В данной статье исследуется применение метода многомерного многошагового прогнозирования на основе модели глубокого обучения LSTM для временных рядов цен биткойнов и оценивается его эффективность. Проанализированы варианты реализации многомерного многошагового прогнозирования на основе глубокого обучения LSTM и выбран прямой подход к построению многошаговых прогнозов.

Ключевые слова: прогнозирование, машинное обучение, LSTM, Биткойн, кумулятивная стабильность, просадки.

Введение

В последние годы значительно расширилось применение криптовалюты для различных торговых операций, для инвестирования. Во многих странах приняты юридические нормы для ее регулирования и использования в расчетах. Современная ситуация, сложившаяся в финансовом секторе, способствует и в дальнейшем существенному повышению количества операций с криптовалютой, усилению ее правового признания. Это способствует снижению непредсказуемой волатильности цен на криптовалюту и ее волатильность немного снижается. А, значит, может повышаться эффективность процессов прогнозирования и качества получаемых прогнозов на цены криптовалюты. Основной криптовалютой остается биткойн и его значение повышается.

Как известно, биткойн, BTC – это децентрализованная цифровая валюта, которая использует технологию блокчейна, то есть это виртуальная валюта [1]. Существует множество различных факторов, которые могут повлиять на цену биткойна. Среди таких факторов можно выделить: общие факторы, влияющие на все рынки, макропоказатели развития экономики, информация о широком внедрении новых технологий, значительные социальные и политические события и прочее. Если имеется неопределённость ситуации, то цены на биткойн могут и снижаться. В настоящее время большинство экспертов в целом настроены оптимистично в отношении цены биткойна и считают, что она будет расти (<https://digitalcoinprice.com>). Однако существуют и сомнения в долгосрочной перспективе роста биткойна, даже не взирая на то, что рынок криптовалют постоянно развивается: все больше и больше компаний внедряют продукты и услуги, связанные с криптовалютой (<https://walletinvestor.com>). Осторожные прогнозы можно объяснить страхом и непониманием технологий криптовалют, а также достаточно жесткими правилами и юридическими нормами их обращения. Поэтому инвесторы с осторожностью инвестируют в криптовалюты. Для повышения уверенности инвесторов и трейдеров в инвестировании в криптовалюту необходима разработка и совершенствование алгоритмов прогнозирования цен криптовалюты и, в частности цен на биткойн.

Известно, что первые исследования по прогнозированию цен на биткойн относятся к концу 2017 году, хотя биткойн появился в 2008 году. Процесс прогнозирования цен на биткойн относится к достаточно сложным прогнозам временных рядов поскольку его цена характеризуется достаточно большой волатильностью и колебаниями цен за год. Но не взирая на эту волатильность разрабатываются новые подходы и модели для прогнозирования цен на биткойн [2-5]. Основным недостатком текущих исследований по применению моделей машинного обучения для задач прогнозирования временных рядов финансовых активов является тот факт, что прогнозы делаются на короткий период, например день или два. В этом исследовании мы намерены сделать более сложный 7-дневный прогноз, так как это даст нам представление о краткосрочном движении актива. Более того, это позволит нам протестировать модели в более сложной постановке задачи. Поэтому в исследовании будут рассмотрены новые параметры и подходы к задаче прогнозирования.

Целью данной работы является исследование практической реализации метода многомерного многоэтапного прогнозирования на основе модели глубокого обучения LSTM для временных рядов цены на биткойн и оценка эффективности прогнозных данных. Объектом исследования является прогнозирование временных рядов. Предметом исследования является многомерное многоэтапное прогнозирование временного ряда цены Bitcoin на 7 дней вперед с использованием модели глубокого обучения LSTM.

1. Литературный обзор

В результате крайне важно оценить множество моделей и стратегий, прежде чем выбрать лучшую. Многие методы подходят для прогнозирования цен на биткойны, и было проведено много исследований по прогнозированию цен на акции. Криптографические активы, наоборот, требуют более сложных и сложных методов из-за их высокой волатильности. Важно отметить, что в истории фондовой биржи прогнозирование всегда было основным приоритетом. В результате значительное количество ученых создали и обосновали конкретные методы прогнозирования будущих цен на различные активы [5-10].

Поворотным моментом в этой области стало создание авторегрессионного интегрированного скользящего среднего (ARIMA) Боксом и Янкинсом в 1970 году. Даже сейчас эта модель активно используется в гибридных или ансамблевых моделях, а ARIMA является основным эталоном для измерения времени. серийная регрессия (Броквелл П.Дж., Дэвис Р.А., 2009) [11].

В целом, во многих исследованиях для решения этой задачи применялись подходы машинного обучения. Так, в исследовании Окаши (2014) сообщалось, что SVM (Support Vector Machine) способна превосходить ARIMA и даже простые нейронные сети [12]. Что касается нейронных сетей, то в 2015 году Чен К. и Чжоу Ю. применили длинную краткосрочную память для прогнозирования фондового рынка Китая, и было упомянуто, что она имеет отличный потенциал в задачах последовательности, таких как прогнозирование цен акций [13]. В наши дни в широком спектре исследований используются гибридные и сложные модели, основанные на LSTM, для прогнозирования будущих цен. Например, Ву С. 2022 создал модель LSTM на основе сентиментального анализа, данных OHLC и технических индикаторов для прогнозирования цен на фондовом рынке на фондовом рынке Китая [14]. Автор проделал колоссальный объем работы в ходе расследования, создав технические индикаторы для индексации временных рядов и анализа настроений с использованием заголовков новостей, затем создал несколько фреймов данных для источников данных, включая все данные для временных рядов, и спрогнозировал будущие цены. В итоге несколько источников данных продемонстрировали лучшие результаты по метрикам MAE, но по MSE они оказались вторыми.

Недавние исследования были сосредоточены на передовых методах на основе LSTM и гибридных моделях, которые используют различные источники данных, включая анализ настроений, данные OHLC и технические индикаторы [2, 3, 5]. Эти разработки повысили точность прогнозов, продемонстрировав перспективность машинного обучения и анализа временных рядов для прогнозирования цен финансовых активов.

2. Методология

Одномерные и многомерные временные ряды. Одномерный временной ряд – это ряд с одной переменной, которая зависит только от времени. Простейшим примером такого ряда может быть временной ряд значений температуры окружающей среды за каждый час на протяжении нескольких лет, например за последние три года. Такой ряд характеризуется тем, что в нем температура, как зависимая переменная, определяется только временем. В данном случае можно анализировать только одну переменную – температуру по прошлым историческим показателям. При этом можно находить различные закономерности, например, можно выделить сезонность данных (ночь – день), цикличность (лето-зима). Такой анализ называется одномерным анализом временных рядов.

Если временной ряд имеет больше одной переменной, то это многомерный временной ряд. При этом каждая переменная зависит не только от своих прошлых значений, но также имеет некоторую зависимость от других переменных ряда. Данная зависимость используется для анализа временного ряда и прогнозирования будущих значений. Таким образом, многомерный временной ряд имеет более чем одну переменную, зависящую от времени, и, в свою очередь, каждая из этих переменных определяется не только своими прошлыми значениями, но также имеет некоторую зависимость от других переменных, входящих в данный многомерный ряд.

При анализе многомерных временных рядов используют зависимости между рядами для обеспечения более надежных и точных данных анализа и прогнозирования. Устанавливается зависимость между рядами для того, чтобы определить, полезен ли конкретный временной ряд, например, Y_2 , для прогнозирования другого ряда, например, Y_1 . Для этой цели используется тест причинности Грейнджера, предложенный автором в 1969 году [16]. Автор теста основывался на том факте, что причинно-следственная связь может быть проверена путем измерения способности предсказывать будущие значения данного временного ряда, используя предыдущие значения другого временного ряда. В данном случае не проверяется истинная причинно-следственная связь между двумя рядами, поскольку факт коррелированности данных во многих случаях не обязательно обоснован причинно-следственной связью. Достаточно наглядно данный эффект можно увидеть визуально на графиках рядов: когда Y_2 следует некоторым шаблонам, которые имеет Y_1 после некоторой временной задержки. Тест Грейнджера находит только «прогностическую причинность». Таким образом, проводится проверка статистической гипотезы для определения того, полезен ли один временной ряд для прогнозирования другого (с помощью серии t-тестов и F-тестов). При этом нулевая гипотеза состоит в том, что запаздывающие значения Y_2 не объясняют изменение Y_1 , поэтому Y_2 не объясняет Y_1 . При критическом значении 5%, если значение p , полученное в результате теста, меньше уровня значимости 0,05, то можно отклонить нулевую гипотезу и с уверенностью 95% установить, что Y_2 можно использовать для прогнозирования Y_1 . Значение p измеряет вероятность того, что любое наблюдаемое различие между группами обусловлено случайностью. Применение данного теста возможно только для стационарных временных рядов. Если какой-либо из рядов не является стационарным, то его необходимо сначала сделать стационарным с помощью соответствующего преобразования.

3. Практические результаты

Многомерный многошаговый метод прогнозирования. Был исследован многомерный многошаговый метод прогнозирования на примере применения модели машинного обучения *LSTM* для прогнозирования курса биткойна на 7 дней вперед. Модель *LSTM* отличается от всех остальных в отношении многомерности, поскольку она может принимать на вход несколько временных рядов и сразу же выдавать на выходе многошаговое предсказание. Благодаря своей способности распознавать сложные временные корреляции и закономерности сети *LSTM* продемонстрировали эффективность в задачах многоэтапного прогнозирования временных рядов. Модель *LSTM*, например, можно обучить, используя ряд исторических ценовых входных данных для прогнозирования цены Биткойна. Скользящие окна или методы последовательности за последовательностью (*seq2seq*) могут использоваться для адаптации модели *LSTM* к созданию многошаговых прогнозов цены биткойна. Для прогнозирования использовался метод прямого прогноза. Формируется 7 моделей для многошаговых прогнозов на 7 дней. Каждая модель эффективно предсказывает свой собственный период. После этого происходит объединение всех результатов и получается прогноз на 7 дней вперед. Этот подход требует больших вычислительных мощностей. Сеть двунаправленной долговременной кратковременной памяти (*LSTM*) для данного исследования была разработана для исследования в бакалаврской работе Саидом Дандамаевым с использованием библиотеки Keras.

Данные взяты за период 2014-2023г.г. (16.06.2023), при этом за период 01.11.2022-31.05.2023 – тестовая выборка. Целевая цена биткойна в данном исследовании – это цена закрытия, которую и прогнозируем.

Для прогнозирования использовались три модели: в одной модели – одномерной использовались только показатели *OHLCV*, в двух других моделях (многомерных) использовались два временных ряда данных: в одном *OHLCV* и временной ряд кумулятивного показателя стабильности цен биткойна, а в другом – *OHLCV* и временной ряд кумулятивного показателя просадок цен биткойна. Дополнительным временным рядом в модели многомерного прогнозирования является динамический ряд кумулятивных показателей стабильности исходного временного ряда котировок биткойна. Динамический ряд кумулятивных индикаторов стабильности характеризует тенденцию относительно стабильности уровней временного ряда котировок биткойна (более полное описание приведено в главе 2 монографии). Стабильность определяется как значение уровня временного ряда, равное или превышающее предыдущие уровни. Этот кумулятивный подход к оценке стабильности цены биткойна позволяет нам учитывать рост/падение цены и величину этого роста/падения. Совокупные показатели стабильности временных рядов цен на биткойн можно рассматривать как информативный индикатор качества поведения цены биткойна на фондовом рынке. Этот индикатор сочетает в себе как рисковую,

так и ценовую составляющие. Косвенно кумулятивные показатели стабильности также учитывают и просадки. Графики временных рядов для анализируемых цен биткоина показаны на рис. 1.



Рис. 1. Временные ряды кумулятивного показателя стабильности и кумулятивной просадки цен биткоина

Архитектура модели и гиперпараметры. Сеть двунаправленной долговременной краткосрочной памяти (LSTM) была разработана с использованием библиотеки Keras. Мы выбрали библиотеку Keras, потому что она имеет удобный интерфейс и синтаксис верхнего уровня для реализации этих моделей. Более того, он обладает обширной функциональностью, позволяющей усложнять модели и улучшать процесс обучения (например, раннюю остановку и сокращение обратных вызовов lr). Архитектура выглядит следующим образом:

Двунаправленный уровень LSTM с 200 единицами и регуляризацией ядра L2 ($\lambda=0,02$).

Слой Dropout с коэффициентом исключения 0,4 для предотвращения переобучения.

Слой LSTM со 150 единицами и регуляризацией ядра L2 ($\lambda=0,01$).

Слой Dropout с коэффициентом исключения 0,2.

Плотный слой с 48 единицами и функцией активации ReLU для нелинейного преобразования.

Слой Dropout с коэффициентом исключения 0,45.

Выходной плотный слой с 1 единицей, прогнозирующий будущую цену биткойнов.

Бэкестирование. Процесс проверки модели будет структурирован следующим образом. У нас есть исторические данные о ценах на биткойны с сентября 2014 года по 16 июня 2023 года. Однако нам также нужна проверочная выборка, которая будет использоваться для выбора количества эпох и гиперпараметров. В качестве проверочной выборки (для выбора гиперпараметров модели) было решено выбрать интервал с марта 2022 г. по октябрь 2022 г., а для финального тестирования — интервал с ноября 2022 г. по июнь 2023 г.

Тестирование будет проходить без предварительного обучения — т. е. модель будет обучена только на данных до марта 2022 года, а затем протестирована на остальных данных.

Что касается итогового тестирования, то мы разделим тестовую выборку на необходимое количество частей и затем замерим метрики на каждой части. После этого мы измерим среднее значение показателей на каждом сгибе и запишем это в окончательный результат. Стоит отметить, что мы также будем измерять качество моделей при прогнозировании всех шагов от 1 до 7. Таким образом, мы будем тестировать каждую модель при прогнозировании на разное количество дней, а не только на 7.

Составление прогноза. Для составления прогноза мы использовали метод прямого прогнозирования. Этот метод работает по следующему принципу: для создания прогноза в момент времени $t+1$ мы обучаем первую модель, которая учится предсказывать значения в момент времени $t+1$ на основе текущих значений. Затем мы строим вторую модель, которая учится прогнозировать значения в момент времени $t+2$, используя текущие значения. В результате у нас будет n моделей многошаговых прогнозов на n дней вперед. В нашем случае это будет 7 моделей. В ходе обучения мы обеспечим эффективность каждой модели при прогнозировании своего периода. После этого будет финальная валидация, где мы объединим все прогнозы и получим прогноз на 7 дней вперед. Этот подход требует большой вычислительной мощности, поскольку мы обучаем 7 моделей LSTM для каждого набора данных. Это необходимо учитывать при возможном использовании этого подхода в будущем.

Результаты экспериментов и анализ. Что касается метрики по индексу стабильности, то следует подчеркнуть, что временной ряд Биткойн имеет высокий показатель индекса стабильности, 87,3% (рассчитывается как отношение суммы показателей стабильности временного ряда к сумме показателей стабильности полностью стабильного ряда за рассматриваемый временной период). Можно видеть, что в периоды бычьего рынка цены биткойнов достигают новых верхних значений.

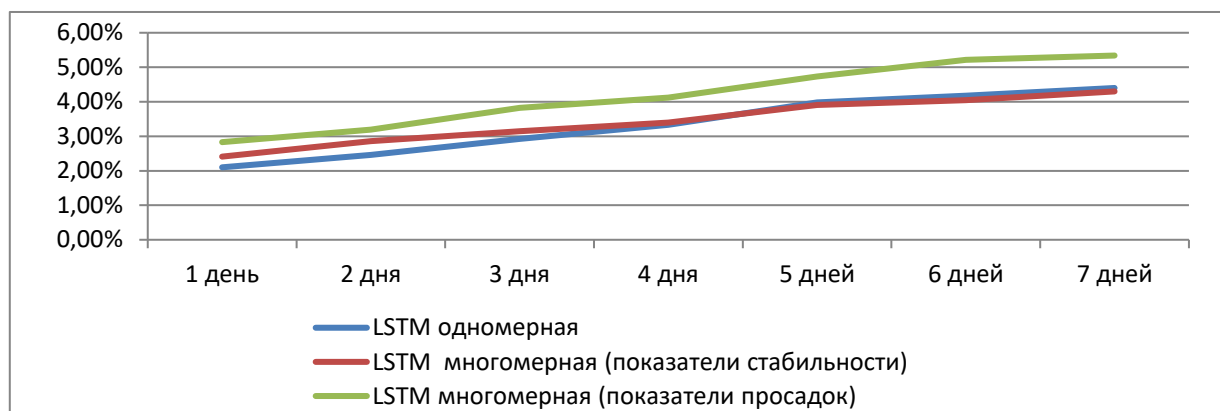


Рис. 2. Метрика MAE по прогнозированию цены закрытия биткойна по модели LSTM на июнь 2023г.

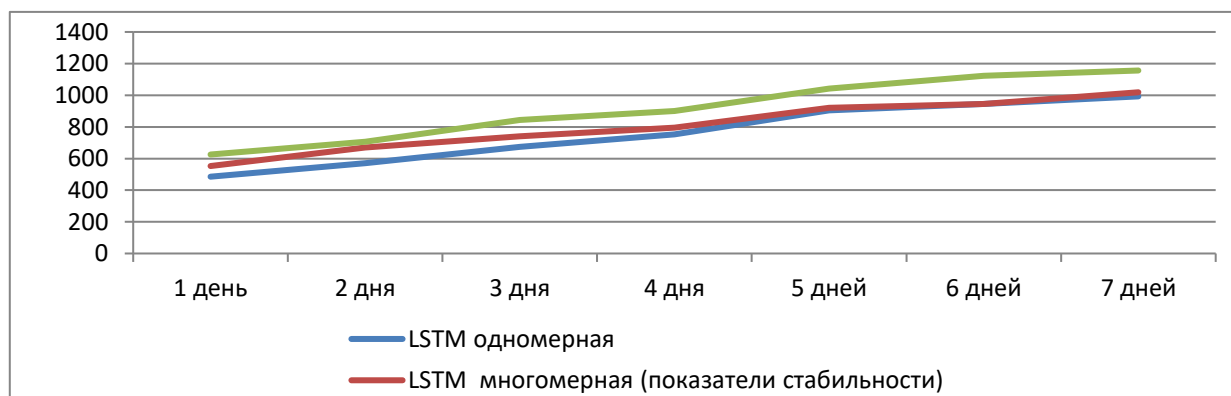


Рис. 3. Метрика SMAPE по прогнозированию цены закрытия биткойна по модели LSTM на июнь 2023г.

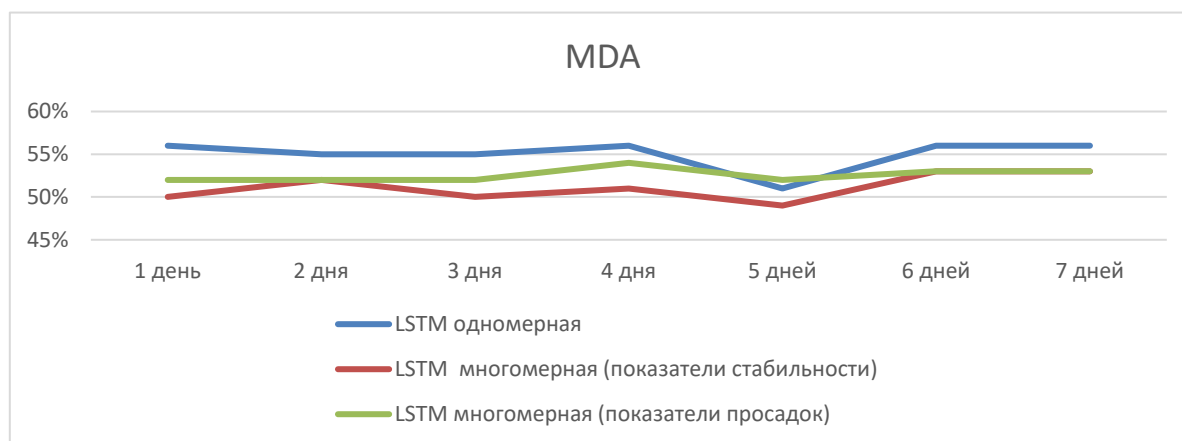


Рис. 4. Метрика MDA по прогнозированию цены закрытия биткойна по модели LSTM на июнь 2023г.

Выходной переменной, как результатом прогнозирования, является цена закрытия торгового дня по биткойну. Результаты показывают, что при небольшом горизонте прогнозирования модель, обученная на торговых данных, показывает наилучшие результаты. В то же время модель, обученная на данных с функцией стабильности, показала несколько лучшие результаты по процентной ошибке. Следует отметить, что временной ряд показателей стабильности достаточно сильно коррелирует с временным рядом показателей OHLCV, но он отражает показатели отклонений от стабильности. Это позволяет уточнить и скорректировать прогнозные показатели. Что касается временного ряда кумулятивных

просадок, то этот показатель ухудшил результаты модели, из чего можно сделать вывод о его незначительности влияния на прогностическую способность модели. По-видимому, он будет хорошо работать в случае нестабильного временного ряда цен биткоина.

Подобранная архитектура многомерной многоэтапной модели позволила повысить точность прогноза по сравнению с одноэтапными моделями на 10-15%. Применение многомерных многоэтапных моделей может позволить получить более точный прогноз. В нашем случае многомерная многоэтапная модель с дополнительным временным рядом кумулятивных показателей устойчивости позволила получить лучшие результаты на 5 и более шагов прогнозирования. Это показывает, что необходимо дальнейшее исследование многомерных многоэтапных моделей прогнозирования с дополнительными данными, например техническими индикаторами, дополнительными финансовыми показателями, например, временными рядами цен на золото, нефть и прочее.

4. Заключение

Метод прямого многошагового прогнозирования имеет потенциал в контексте сетей LSTM. В то же время потенциал развития таких моделей достаточно обширен. Например, каждую одношаговую модель можно улучшить, используя родственные архитектуры (CNN-LSTM, Encoder-Decoder LSTM и другие), а также гибридные подходы (например, разложение временного ряда на дорожку и шумовую составляющую и затем применив несколько моделей, включая SVR, ETS и LSTM). Также можно рассмотреть более обширный поиск по сетке, чтобы найти наилучшее количество слоев и гиперпараметров.

Стоит отметить, что мы провели обширный анализ, построение и обучение многих моделей. Однако прямой подход считается ресурсоемким, а обучение нейронных сетей требует графических ресурсов. Более того, хотя наши ресурсы позволяют нам проводить обширный анализ, различные подходы глубокого обучения и гибридные модели требуют для анализа большого количества вычислительных ресурсов. Таким образом, подходы, которые мы перечислили выше, могут повысить точность прогнозирования, если будет предоставлено достаточно ресурсов.

Такая выбранная архитектура многоэтапной модели позволила повысить точность прогнозирования по сравнению с одноэтапными моделями на 10-15%.

Применение многомерных многошаговых моделей может позволить получить более точные прогнозы. В нашем случае многомерная многошаговая модель с разными временными рядами совокупных показателей устойчивости показала лучшие результаты для пяти и более шагов прогнозирования. Это показывает, что необходимо дальнейшее исследование многомерных моделей многошагового прогнозирования с дополнительными данными, такими как технико-экономические показатели, такие как цены на золото и нефть.

Литература

1. *Harwick Cameron*. Cryptocurrency and the Problem of Intermediation. Independent Review, Vol. 20, No. 4, Spring 2016.
2. *Calzone, O.* (2023, June 30). An intuitive explanation of LSTM. Medium. <https://medium.com/@ottavioalzone/an-intuitive-explanation-of-lstm-a035eb6ab42c>.
3. *Li, Y., Jiang, S., Li, X., & Wang, S.* (2022). Hybrid data decomposition-based deep learning for bitcoin prediction and algorithm trading. Financial Innovation, 8(1), 1-24.
4. *Zi Ye, Yin Xu Wu, Hui Chen, Yi Pan, Qingshan Jiang.* A Stacking Ensemble Deep Learning Model for Bitcoin Price Prediction Using Twitter Comments on Bitcoin. Mathematics 2022, 10(8), 1307; <https://doi.org/10.3390/math10081307>.
5. *Siva Kumar.A;Gopirajan Pv;Beulah Jackson.* Machine Learning-Based Timeseries Analysis for Cryptocurrency Price Prediction: A Systematic Review and Research, 2023,Page(s):1 – 5.
6. *Jaquart, P., Dann, D., & Weinhardt, C.* (2021). Short-term bitcoin market prediction via machine learning. The journal of finance and data science, 7, 45-66.
7. *Guo, H.Z.; Zhang, D.; Liu, S.Y.; Wang, L.* Bitcoin price forecasting: A perspective of underlying blockchain transactions. Decis. Support Syst. 2021, 151, 113650.
8. *Loginova, E.; Tsang, W.K.; van Heijningen, G.; Kerkhove, L.; Benoit, D.F.* Forecasting directional bitcoin price returns using aspect-based sentiment analysis on online text data. Mach. Learn. 2021.
9. *Li, Y.; Dai, W.* Bitcoin price forecasting method based on CNN-LSTM hybrid neural network model. J. Eng. 2020, 2020, 344–347.
10. *Raju, S.M.; Tarif, A.M.* Real-time prediction of BITCOIN price using machine learning techniques and public sentiment analysis. arXiv 2020, arXiv:2006.14473.

11. *Brockwell P. J., Davis R. A.* Time series: theory and methods. – Springer science & business media, 2009.
12. *Okasha M. K.* Using support vector machines in financial time series forecasting //International Journal of Statistics and Applications. – 2014. – T. 4.– №. 1. – pp. 28-39.
13. *Chen K., Zhou Y., Dai F.* A LSTM-based method for stock returns prediction: A case study of China stock market //2015 IEEE international conference on big data (big data). – IEEE, 2015. – C. 2823-2824.
14. *Wu S. et al.* S_I_LSTM: stock price prediction based on multiple data sources and sentiment analysis //Connection Science. – 2022. – T. 34. – №. 1. – pp. 44-62.
15. *Hochreiter, S., & Schmidhuber, J.* (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
16. *Michałków J., Sakowski P., Ślepaczuk R.* LSTM in Algorithmic Investment Strategies on BTC and S&P500 Index //Sensors. – 2022. – T. 22. – №. 3. – p.917.
17. *Nazir S., Azlan Ab Azizl , J. Hosenl , Nor Azlina Azizl , G. Ramana Murthy.* Forecast Energy Consumption Time-Series Dataset using Multistep LSTM Models To cite this article: S. Nazir et al 2021 J. Phys.: Conf. Ser. 1933 012054.
18. *Laptev N., Yosinski J., Li Erran Li, Smyl S.* Time-series Extreme Event Forecasting with Neural Networks at Uber. ICML 2017 Time Series Workshop, Sydney, Australia, 2017. ile:///C:/Users/snv/Downloads/TSW2017_paper_3.pdf.
19. *Yunpeng L., C. Software, and X. Jiaotonguniversity,* “Multistep ahead time series forecasting for different data patterns based on LSTM recurrent neural network,” pp. 305–310, 2017, doi: 10.1109/WISA.2017.25.
20. *Xiaorui Shao, Chang Soo Kim.* Multi-Step Short-Term Power Consumption Forecasting Using Multi-Channel LSTM With Time Location Considering Customer Behavior. *Journals & Magazines, IEEE Access*, Volume: 8.
21. *Sizykh, D., Sizykh, N.* Classification Analysis Method for Dynamic Characteristics of IT Companies Shares Exchange Value. *Proceedings of 2018 11th International Conference " Management of Large-Scale System Development", MLSD 2018*, 2018, <https://ieeexplore.ieee.org/document/8551785>.