**ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ**

**ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ**

**«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

Факультет компьютерных наук Образовательная программа «Программная инженерия»

УДК 004.002, 004.005, 004.062

СОГЛАСОВАНО

Старший преподаватель ФКН

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Д.В.Пантюхин

«01» февраля 2022 г.

УТВЕРЖДАЮ

Академический руководитель образовательной программы

«Программная инженерия», профессор департамента программной инженерии, канд. техн. наук

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_В. В. Шилов

«\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_2022 г

ОТЧЕТ О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

Исследование алгоритмов трейдинга криптовалют с использованием нейросетей

по направлению подготовки бакалавров 09.03.04 «Программная инженерия»

СПИСОК ИСПОЛНИТЕЛЕЙ

Студент образовательной программы 09.03.04

«Программная инженерия»

Мелехин Денис Антонович

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2022.02.01

**Москва 2022**

# РЕФЕРАТ

**Ключевые слова:** *нейронные сети, криптовалюта, блокчейн, предсказание, временные ряды, обучение с подкреплением, большие данные, трейдинг.*

## Оценка современного состояния решаемой проблемы:

Многие трейдеры по всему миру используют различные алгоритмы машинного обучения для торговли криптовалютой. Данное исследование направлено на сравнение существующих решений алгоритмической торговли с новыми алгоритмами.

## Актуальность:

Принятие решения о покупке/продаже определенных валют давно сопровождаются предсказаниями алгоритмов для минимизации рисков. Алгоритмы машинного обучения позволяют трейдерам более грамотно принимать решения. Но в случае дневной и high frequency торговли человек просто не способен настолько быстро анализировать рынок, поэтому для краткосрочной торговли используются только определенные алгоритмы.

## Объект разработки:

Алгоритм машинного обучения для торговли криптовалютой.

Цель исследования:

- Исследовать существующие подходы алгоритмической торговли

- Выявить сильные и слабые стороны существующих решений

- Создать свои модели торговли криптовалютой

- Протестировать созданные модели

- Сравнить созданные модели с существующими решениями

## Задачи исследования:

- Создать программу для получения котировок с биржи и торговли

- На полученных с биржи котировках протестировать существующие решения на разных промежутках времени

- Написать свои модели торговли

- Протестировать модели на написанной программе

## Методы исследования:

* Изучение зарубежных статей по необходимым структурам данных, алгоритмам;
* Выведение формул для понятия принципа работы алгоритмов;
* Обработка входных данных с помощью машинного обучения.

## 

## Результаты работы:

Обученный алгоритм машинного обучения для торговли криптовалютой.

## Область применения:

Трейдинг криптовалютой.

## Практическая значимость:

Обученные алгоритмы машинного обучения позволят выявлять слабые места на рынке криптовалют и автоматически проводить сделки.

## Прогнозные предположения о развитии проекта исследования:

При успешном получении ожидаемых результатов возможно использование алгоритма трейдерами.

**СОДЕРЖАНИЕ**

[РЕФЕРАТ 2](#_Toc95953515)

[Оценка современного состояния решаемой проблемы: 2](#_Toc95953516)

[Актуальность: 2](#_Toc95953517)

[Объект разработки: 2](#_Toc95953518)

[Задачи исследования: 2](#_Toc95953519)

[Методы исследования: 3](#_Toc95953520)

[Результаты работы: 3](#_Toc95953521)

[Область применения: 3](#_Toc95953522)

[Практическая значимость: 3](#_Toc95953523)

[Прогнозные предположения о развитии проекта исследования: 3](#_Toc95953524)

[ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ 5](#_Toc95953525)

[ВВЕДЕНИЕ 6](#_Toc95953526)

[АНАЛОГИ 7](#_Toc95953527)

[1. Facebook Prophet 7](#_Toc95953528)

[2. Gate.io 7](#_Toc95953529)

[МЕТОДЫ УЛУЧШЕНИЯ ПРОГНОЗИРОВАНИЯ 7](#_Toc95953530)

[ДАННЫЕ 8](#_Toc95953531)

[ОПИСАНИЕ АЛГОРИТМОВ 8](#_Toc95953532)

[Метрики 8](#_Toc95953533)

[Риск-менеджмент 9](#_Toc95953534)

[1. LSTM система 9](#_Toc95953535)

[2. DeepAR 12](#_Toc95953536)

[ЗАКЛЮЧЕНИЕ 12](#_Toc95953537)

[НАПРАВЛЕНИЕ ДАЛЬНЕЙШЕЙ РАБОТЫ 12](#_Toc95953538)

[СПИСОК ИСТОЧНИКОВ 13](#_Toc95953539)

# ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ

В настоящем отчете о научно-исследовательской работе применяются следующие термины с соответствующими определениями (таблица 1).

Таблица 1 – Термины и определения

|  |  |
| --- | --- |
| Термин | Определение |
| Криптовалюта | Разновидность цифровой валюты, учёт внутренних расчётных единиц которой обеспечивает децентрализованная платёжная система (нет внутреннего или внешнего администратора или какого-либо его аналога), работающая в полностью автоматическом режиме |
| Трейдер | Торговец, действующий по собственной инициативе, и стремящийся извлечь прибыль непосредственно из процесса торговли. |
| Машинное обучение | Раздел изучения искусственного интеллекта, который изучает методики построения алгоритмов, способных обучаться посредством тех или иных способов. |

# 

# ВВЕДЕНИЕ

В настоящее время криптовалюта развилась в огромную индустрию. Ежедневно множество людей занимаются куплей, продажей и обменом криптовалют. Сделки совершаются как с использованием только криптовалют, так и с использованием реальных денег. Существует множество криптовалютных бирж, которые позволяют пользователям совершать операции с криптовалютами. К примеру, объем рынка криптовалюты в России, на февраль 2022 года, по данным Bloomberg, оценивается в 16.5 триллионов рублей.

Торговля на криптовалютной биржи во многом схожа с торговлей на сырьевых и валютных биржах, которые существуют уже ни одну сотню лет. Из-за этой схожести большинство техник и методик, применимых на традиционной бирже, так же применимы для торговли криптовалютой.

Криптовалютой ежедневно обмениваются огромное количество людей, ее курс часто предсказуем, потому что любые высказывания властей или популярных лиц напрямую влияют на ее курс. Алгоритмы машинного обучения способны выявлять направление движения курса, потому что обучаются на исторических данных, в которых выявляют похожие на текущую ситуацию закономерности.

У алгоритмической торговли есть множество преимуществ:

1. Удобство

Алгоритмическая торговля устраняет необходимость круглосуточно наблюдать за рынком и тратить время на выставление сделок.

1. У алгоритмов нет эмоций

Множество трейдеров сталкивается с сильным стрессом перед осуществлением очередной сделки. Алгоритмы же способны хладнокровно анализировать рынок и принимать только обоснованные фактами решения.

1. Скорость и точность

Алгоритмы торговли позволяет торговать со скоростью недоступной человеку. Они способны моментально оценить огромный объем данных и принять решение.

# АНАЛОГИ

## Facebook Prophet

Согласно статье Facebook Prophet, был разработан для прогнозирования большого числа различных бизнес-показателей и строит достаточно хорошие default'ные прогнозы. Кроме того, библиотека дает возможность, изменяя человеко-понятные параметры, улучшать прогноз и не требует от аналитиков глубоких знаний устройства предсказательных моделей.

## Gate.io

Данный алгоритм выделяет непопулярные криптовалюты и распределяет свой портфель в зависимости от вероятности роста. Также, в зависимости от поведения рынка он отказывается от некоторых валют в пользу более перспективных.

# МЕТОДЫ УЛУЧШЕНИЯ ПРОГНОЗИРОВАНИЯ

Многие криптовалюты сильно коррелируют между собой. Благодаря этому можно выделять группы криптовалют и принимать решения о покупке/продаже в зависимости от поведения валют одной группы.

# ДАННЫЕ

Для тестирования алгоритмов трейдинга была выбрана платформа Binance, потому что она позволяет торговать криптовалютой через библиотеку, написанную на языке Python. У данной платформы нет существенных задержек, у нее небольшая комиссия и она удобна в использовании.

Binance предоставляет возможность выгрузки всех исторических данных, на которых и будет происходить обучение алгоритмов. Для обучения были выбраны одноминутные котировки BTC/USDT за 2021 год.

# ОПИСАНИЕ АЛГОРИТМОВ

# Метрики

Перед разработкой любого алгоритма машинного обучения необходимо определить метрики для оценивания качества работы.

Если алгоритм имеет архитектуру предсказания курса криптовалюты, то самой простой метрикой для оценки работы алгоритма является MSE (Mean Squared Error).

Изображение выглядит как текст, часы

Автоматически созданное описание

Данную метрику можно применять для оценки предсказанной моделью свечей (квадрат разности предсказанного значения свечи и фактического).

Для более интерпретируемого метода оценки алгоритма также применяется метрика MAE (Mean Absolute Error)

Изображение выглядит как текст

Автоматически созданное описание

Данная метрика поможет оценить то, сколько денег трейдер заработает/потеряет в каждый момент времени кратный длительности свечи.

# Риск-менеджмент

Один из важнейших моментов при создании алгоритмической торговой системы – это управление рисками. Грамотно составленная стратегия управления рисками позволит избежать значительных потерь капитала. В биржевой торговле эта система называется риск-менеджмент.

**Правила риск-менеджмента** должны быть простыми и однозначными, чтобы они надежно управляли алгоритмом торговли.

В стратегии риск-менеджмента должны быть четко описаны следующие риски:

* **Риск на сделку** – это максимальный риск на сделку, допустимый в системе риск-менеджмента. Максимально рекомендованный риск на сделку – 1% от депозита. Такой небольшой риск повышает прибыльность алгоритма и позволяет пользователю контролировать убытки.
* **Риск на день** – максимальный риск потери за один торговый день, который допущен в вашей системе риск-менеджмента.  Рекомендованный риск на день – не более 3% от депозита.

# LSTM система

LSTM системы часто используются для прогнозирования различных временных рядов, а все котировки ценных бумаг являются временным рядом.

Устройство LSTM сети:

1. Входной Embedding слой:

Слой реализован в модуле nn.Embedding библиотеки pytorch. Данный слой необходим для преобразования каждого слова в вектор. Он принимает на вход вектор, размером с словарь, где на месте значения слова в словаре стоит 1, а на остальных местах нули. Выход этого слоя состоит из 100 нейронов (embedding\_dim), в каждом из которых будет некоторое значение. Вектор из этих выходных значений – это новое векторное представление слова в нейросети. Данный вектор получается для каждого слова и затем каждое слово заменяется этим вектором.

1. LSTM слой:

Данный слой является основным в LSTM сети. Он реализован в модуле nn.lstm библиотеки “pytorch”. В программе данный слой собой блок нейронов с рекуррентными связями. Структура нейрона в LSTM слое представлена на схеме

A picture containing clock, game

Description automatically generated

Рисунок [1][[1]](#footnote-1):

Всего таких нейронов 25 в блоке (hidden\_size).

На вход lstm слой получает выход из Embedding слоя размерностью <batch\_size, min\_words\_in\_sentence, embedding\_dim>. В нашем случае это трехмерный вектор размером <80, 25, 100>.

Выход из lstm слоя имеет размерность <batch\_size, min\_words\_in\_sentence, hidden\_size>. В нашем случае это трехмерный вектор размером <80, 25, 25>.

Слой линейного преобразования:

Данный слой необходим для преобразования выхода lstm слоя. Он принимает выход последнего нейрона lstm слоя размерностью min\_words\_in\_sentence, после чего преобразует данный вектор в скаляр.

1. Сигмоидальное преобразование:

Данный слой необходим для представления выходного скаляра из предыдущего слоя в вероятность принадлежности информативному или не информативному тексту. Функция, применяемая к каждому итоговому значению каждого батча:

A close up of a logo

Description automatically generated

1. Процесс обучения сети:

Обучение производилось оптимизатором из модуля torch.optim.Adam. Обучение на входных данных в программе происходило 4 раза (количество эпох). Для подсчета функции ошибки использовалась бинарная кросс-энтропия (модуль nn.BCELoss). Для регулировки скорости использовался планировщик torch.optim.lr\_scheduler.ReduceLROnPlateau.

Адам – это современный алгоритм, используемый в области глубокого обучения, который быстро достигает отличных результатов.

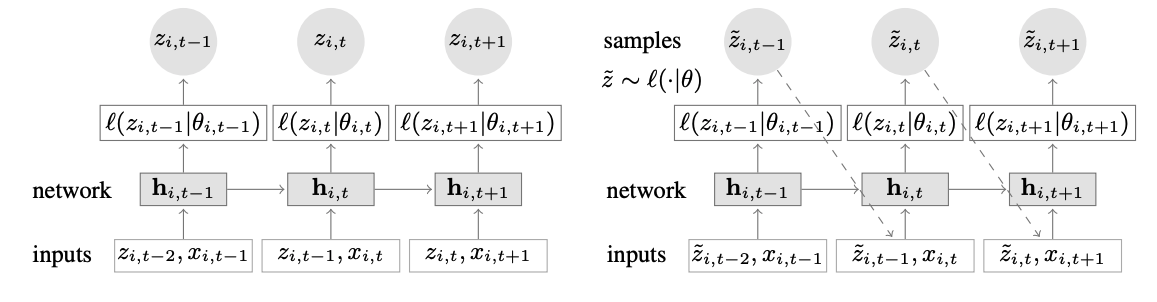
Функция бинарной кросс-энтропии используется для подсчета ошибки в задачах, в которых необходимо распределить входные данные на 2 категории. Данная функция принимает на вход два числа из отрезка [0;1] и возвращает число из этого же отрезка.

Планировщик в нейронной сети необходим для установки более подходящей скорости обучения на для каждой эпохи. Он позволяет нейросети обучаться медленнее в конце обучения.

# DeepAR

Данная модель прогнозирования временных рядов, содержащая в себе элементы глубокого обучения, дает возможность предсказывать вероятности, что позволит трейдеру регулировать риски. Предсказание вероятностей в моделе реализовано методом Монте-Карло.

Архитектура системы:



Рисунок[2] [[2]](#footnote-2):

# ЗАКЛЮЧЕНИЕ

Обученные алгоритмы машинного обучения способны эффективно выявлять слабые места на рынке криптовалют и проводить выгодные сделки в автоматическом режиме. Благодаря настроенному риск-менеджменту система останавливает свою работу в случае превышения допущенного риска. Метрики же позволяют точно определить эффективность модели.

# НАПРАВЛЕНИЕ ДАЛЬНЕЙШЕЙ РАБОТЫ

* Оптимизация алгоритмов машинного обучения
* Добавление новых параметров алгоритмов
* Исследование новых статей

# СПИСОК ИСТОЧНИКОВ

1. https://medium.com/geekculture/what-happens-if-you-day-trade-crypto-with-facebook-prophet-84df66c31657 [Электронный ресурс] / Medium [Официальный сайт]. Режим доступа: свободный (дата обращения: 01.02.22)
2. https://medium.com/@nusfintech.ml/fx-algorithmic-trading-with-machine-learning-models-7d6821fa7e67 [Электронный ресурс] / Medium [Официальный сайт]. Режим доступа: свободный (дата обращения: 01.02.22)
3. https://medium.com/@AlphaGrow/a-machine-learning-framework-for-algorithmic-trading-45575c07c599 [Электронный ресурс] / Medium [Официальный сайт]. Режим доступа: свободный (дата обращения: 01.02.22)
4. https://www.binance.com/en [Электронный ресурс] / Binance [Официальный сайт]. Режим доступа: свободный (дата обращения: 01.02.22)
5. https://python-binance.readthedocs.io/en/latest/ [Электронный ресурс] / Python-Binance docs [Официальный сайт]. Режим доступа: свободный (дата обращения: 01.02.22)
6. https://pandas.pydata.org/ [Электронный ресурс] / Pandas [Официальный сайт]. Режим доступа: свободный (дата обращения: 01.02.22)
7. https://numpy.org/ [Электронный ресурс] / Numpy [Официальный сайт]. Режим доступа: свободный (дата обращения: 01.02.22)

1. Источник рисунка [1]: https://en.wikipedia.org/wiki/Long\_short-term\_memory [↑](#footnote-ref-1)
2. Источник рисунка [2]: https://kshavg.medium.com/deepar-probabilistic-forecasting-with-autoregressive-recurrent-networks-fa65dddae1f5 [↑](#footnote-ref-2)