

Методы оптимизации портфеля криптовалют с учетом новостного фона

Командный курсовой проект

Авторы:

- Барателиа Мирон
- Мищенко Александр

Научный руководитель:

- Мунерман Илья Викторович

Цели проекта

- Исследовать существующие методы оптимизации портфеля криптовалют
- Исследовать способы извлечения и обработки данных о новостном фоне
- Исследовать возможные способы интеграции данных о новостном фоне в различные модели для увеличения точности прогноза

Основные этапы

- Извлечение данных о истории курса криптовалют
- Модель Марковица / Блэка-Литтермана
- Нейронные сети
- Получение новостей
- Недавние нововести (обработка и способы интегрирования)
- Новости за период анализа (обработка и способы интегрирования)
- Улучшение точности анализа новостного фона по новостям
- Представление оптимальной модели, учитывающей новостной фон

Распределение задач

Мирон Барателиа

- Чтение истории криптовалют
- Визуализация данных
- Нейронные сети
- Извлечение новостей и сохранение в таблицу
- Обработка новостей через сентименты
- Совмещение модели Марковица и последних новостей
- Совмещение модели Марковица, нейронной сети и новостного фона.
- Тестирование итоговой модели
- Подготовка отчета и презентации

Александр Мищенко

- Модель Марковица
- Модель Блэка-Литтермана
- Совмещение нейросети и новостного фона
- Подготовка отчета

Хотим заметить, что проект был рассчитан на 3 человека

Извлечение данных

Задача:

Найти способ получения актуальных данных к курсу криптовалют. При этом иметь возможность изменить список исследуемых криптовалют.

Решение:

Запрашиваем данных на API CoinGecko. Это надежный источник, который позволяет бесплатно получить историю многих криптовалют.

Недостатки:

Поскольку источник бесплатный, мы не можем получить точные данные за промежуток более 3 месяцев. Так же не все данные приходят после первого запроса, но это решается повторными запросами.

Визуализация данных

Задача:

Представить полученные данные, для удобного восприятия человека.

Решение:

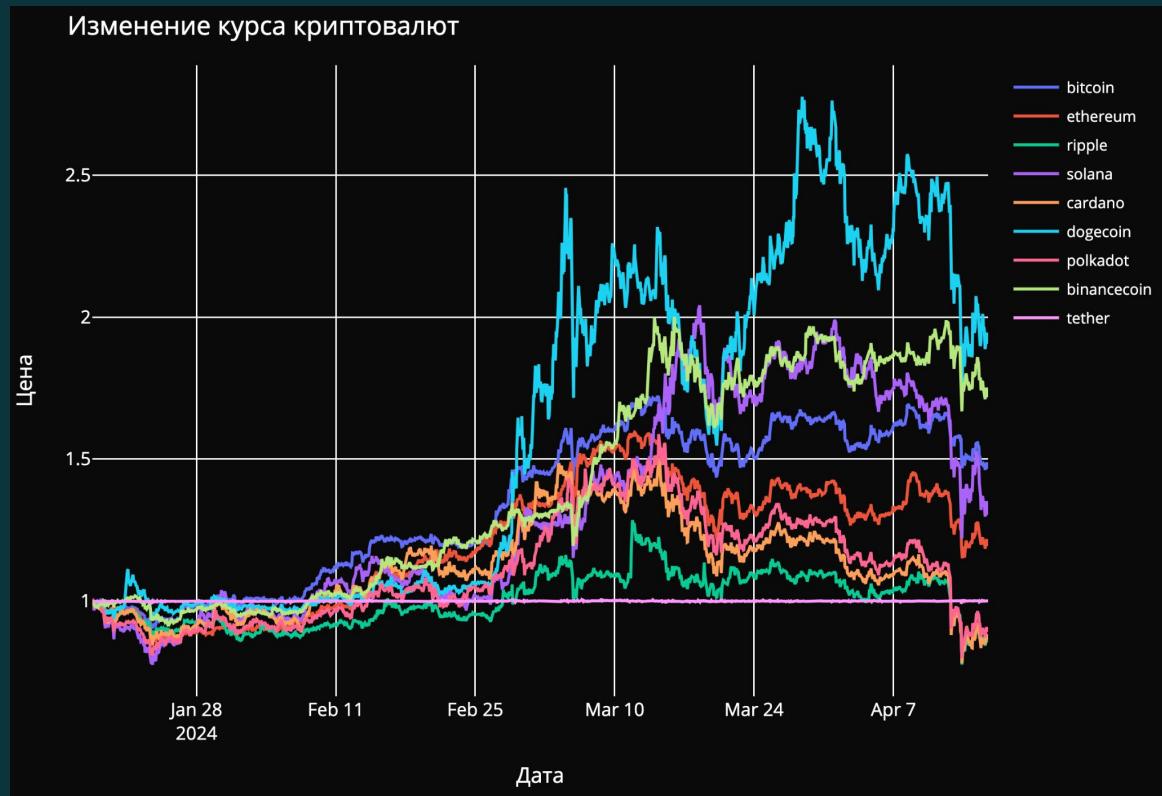
Используем библиотеки matplotlib и plotly для построения графиков.

График изменения курса (для удобства восприятия рассматриваем не саму цену, а только ее изменение, при начальном значении в 1\$).

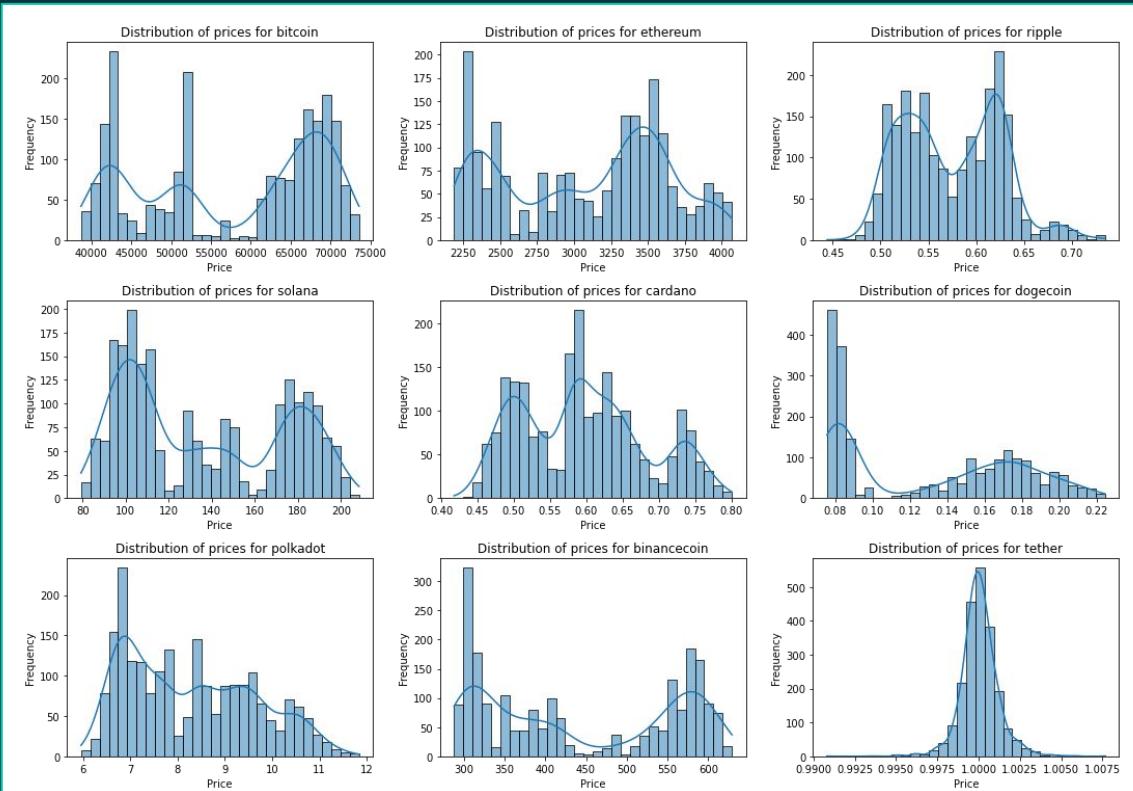
Гистограммы распределения цен

Корреляционная матрица

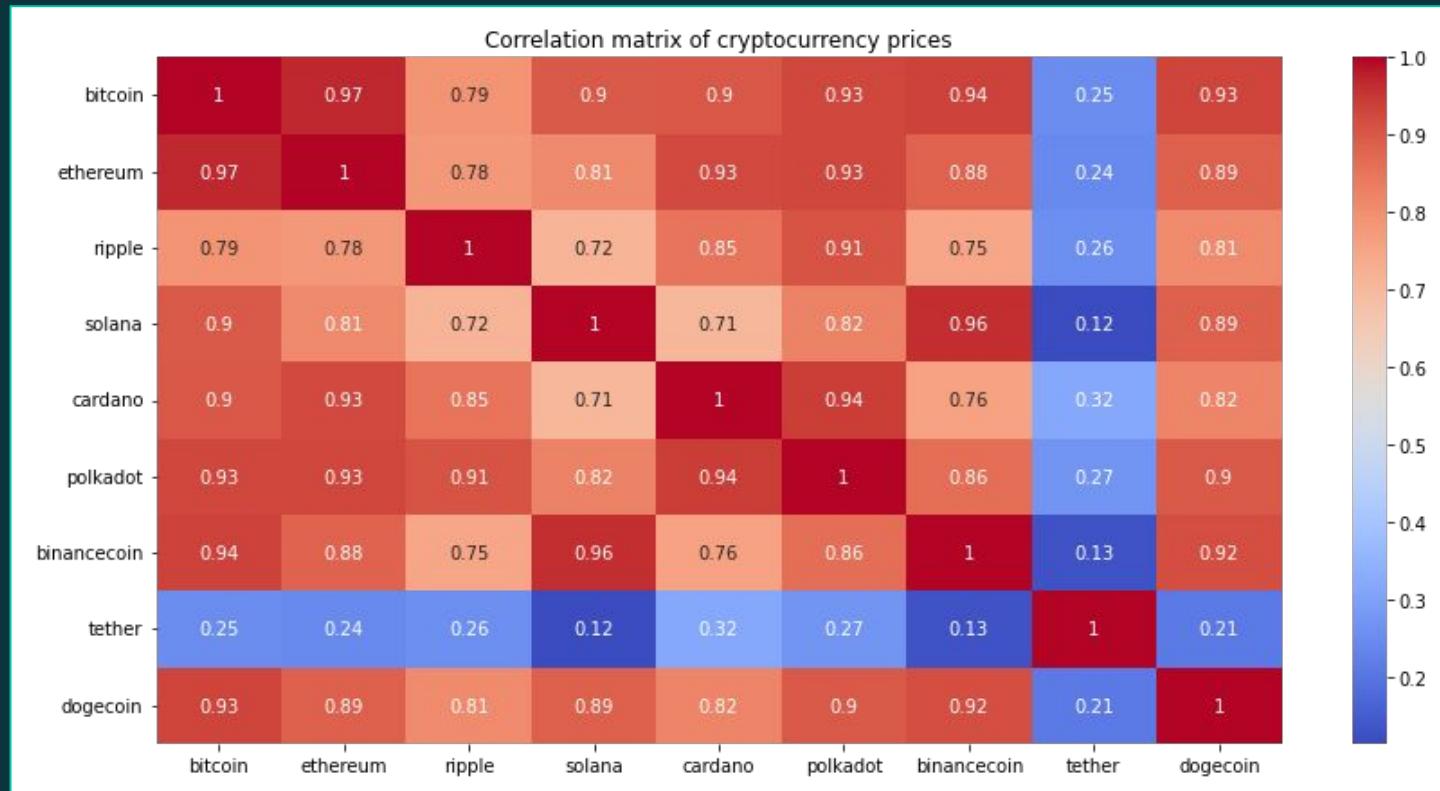
График изменения курса



Гистограммы распределения цен



Корреляционная матрица



Модель Марковица

Описание модели:

Модель Марковица - это теория портфеля, которая позволяет определить оптимальный портфель, минимизируя риск при заданном уровне ожидаемой доходности. Она основана на двух основных параметрах: средней доходности и стандартном отклонении (или волатильности) доходности.

Ожидаемая доходность и волатильность портфеля:

- **Ожидаемая доходность** портфеля рассчитывается как взвешенная сумма ожидаемых доходностей отдельных активов.
- **Волатильность портфеля** (или стандартное отклонение доходности портфеля) рассчитывается как квадратный корень из взвешенной суммы ковариаций доходностей активов.

Модель Блэка-Литтермана

Описание модели:

Модель Блэка-Литтермана - это модификация модели Марковица, которая позволяет инвесторам внести свои собственные прогнозы относительно ожидаемой доходности активов. Это достигается путем введения параметра "доверия" к собственным прогнозам и последующего объединения этих прогнозов с историческими данными.

Отличие от модели Марковица:

Важным отличием модели Блэка-Литтермана от модели Марковица является то, что она позволяет инвесторам учитывать свои собственные прогнозы доходности, а не полагаться только на исторические данные.

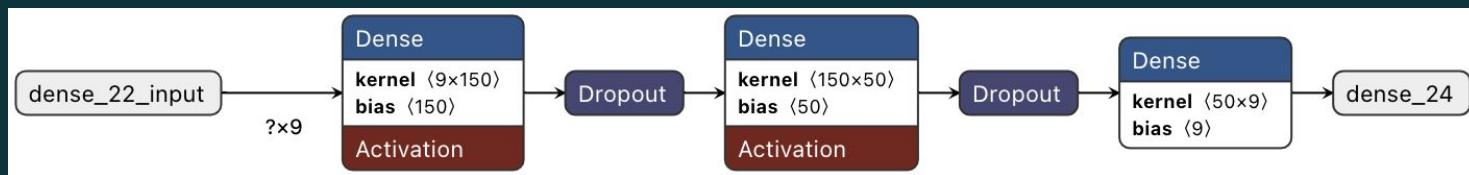
Нейронная сеть

Введение:

Нейронная сеть обучается на исторических данных о доходности криптовалют с целью прогнозирования будущих значений. Затем эти прогнозы используются для определения оптимальных весов портфеля, которые максимизируют ожидаемую доходность.

Метод:

Оптимизация портфеля осуществляется путем минимизации функции `optimize_portfolio`, которая возвращает отрицательное значение прогнозируемой доходности портфеля. Это эквивалентно максимизации прогнозируемой доходности.



Нейронная сеть



Анализ новостей

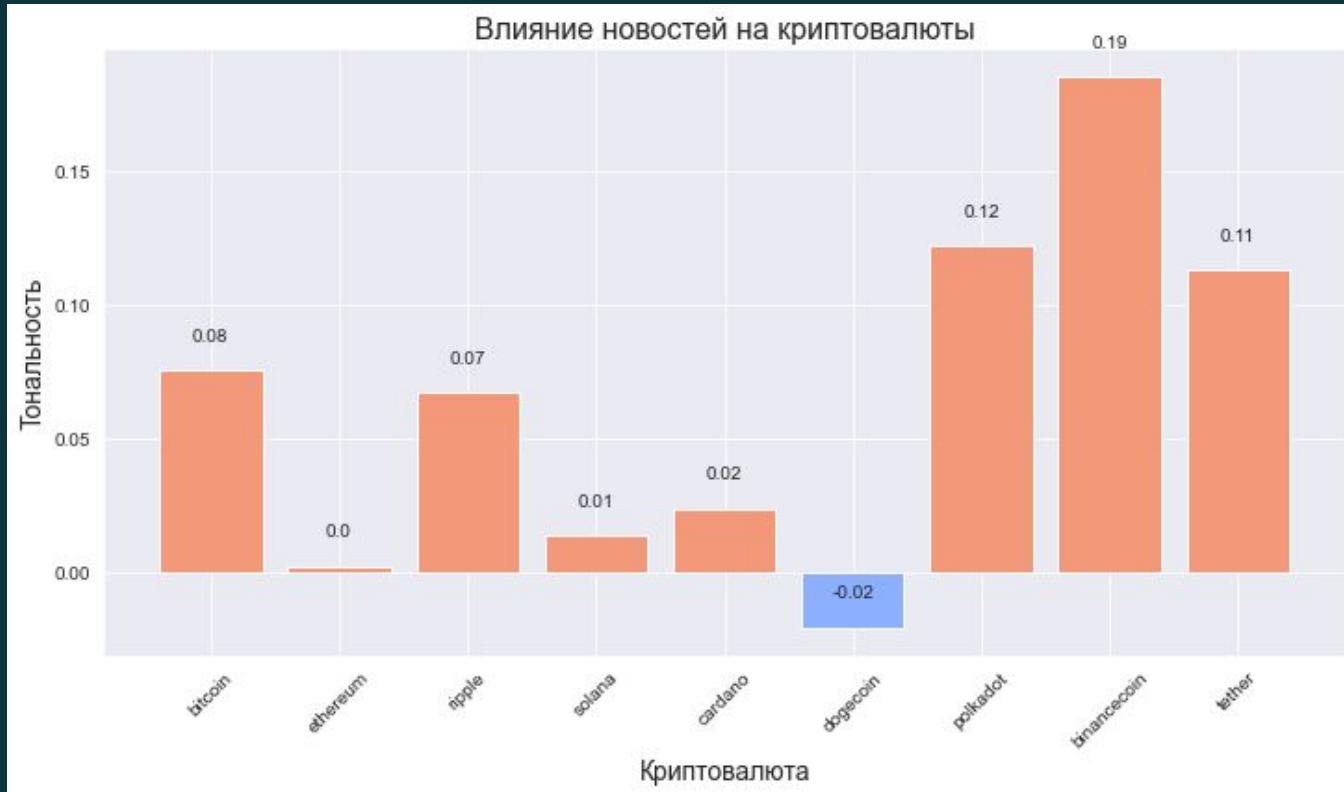
Введение:

В этом разделе мы хотим получить данные о новостях, связанных с криптовалютами, и использовать их для оптимизации портфеля криптовалют. Мы используем API NewsAPI для получения новостей и инструмент SentimentIntensityAnalyzer из библиотеки NLTK для анализа тональности текста новостей.

Анализ новостей:

Мы предполагаем, что эмоциональная окраска новостей может влиять на доходность криптовалюты. Поэтому к средней доходности каждой криптовалюты добавляется соответствующий показатель тональности.

График новостного фона



Обработка новостей за весь промежуток наблюдения

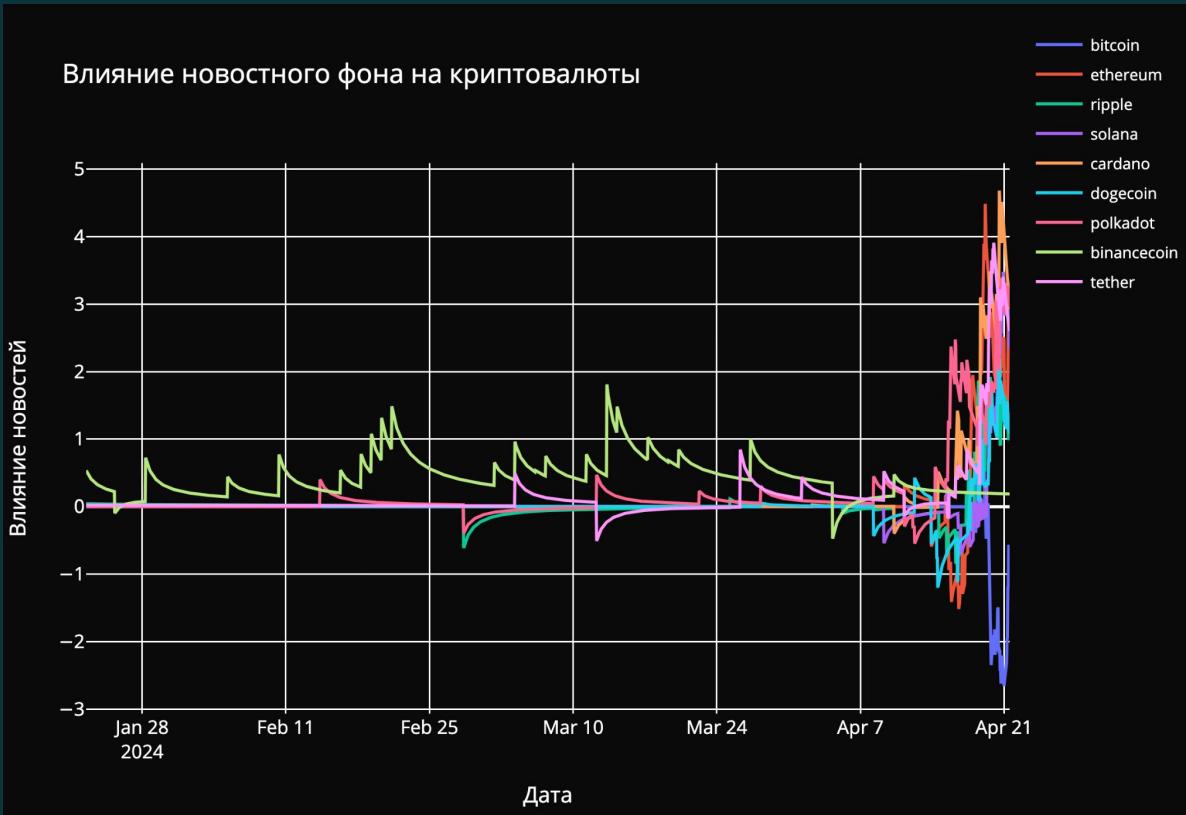
Введение:

В этом разделе мы рассмотрим процесс получения и обработки новостей о криптовалютах за весь промежуток времени. Мы изменили источник данных на Google Search, поскольку источник NewsAPI не предоставляет информацию о старых новостях.

Расчет влияния новостей:

Для расчета влияния новостей на криптовалюты мы используем анализатор настроений. Для каждой новости мы вычисляем ее сентимент и умножаем его на вес, который обратно пропорционален времени, прошедшему с момента публикации новости.

График всех новостей



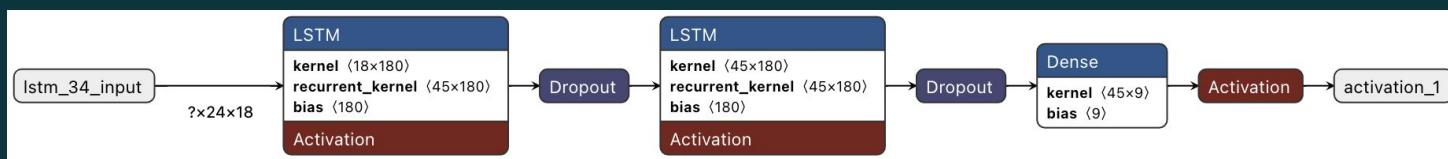
Нейросеть + новостной фон

Введение:

В этом разделе мы рассмотрим улучшение нашей модели прогнозирования доходности криптовалют, добавив в нее новостной фон. Мы используем модель LSTM (Long Short-Term Memory), специально разработанную для работы с временными рядами, и подготавливаем наши данные, используя процентное изменение цен криптовалют.

Обучающая выборка:

Для обучения модели мы используем данные за последние 24 часа, чтобы модель могла улавливать ежедневные тренды и изменения на рынке криптовалют.



Итоговая модель

Введение:

В итоговой модели мы решили скомбинировать результаты нейросети и модель Марковица. Модель Марковица хорошо подходит для оптимизации портфеля, а нейросеть эффективно справляется с обработкой новостей.

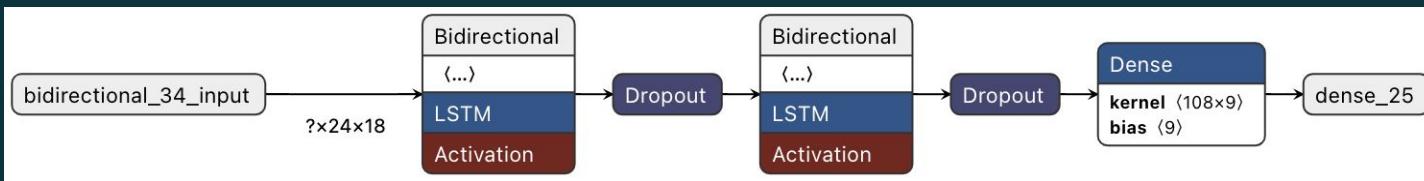
Способ объединения:

Для того чтобы модель Марковица могла учитывать результат нейросети, мы предсказываем не оптимальный портфель, а то, на сколько изменится курс криптовалют.

Итоговая модель

Введение:

1. В качестве целевого значения для нейросети используем процентное изменение курса криптовалют.
2. Обучаем модель



3. Получаем предсказание на изменение курса в ближайший час
4. Находим оптимальный портфель с помощью метода Марковица, опираясь на полученное предсказание

Тестирование

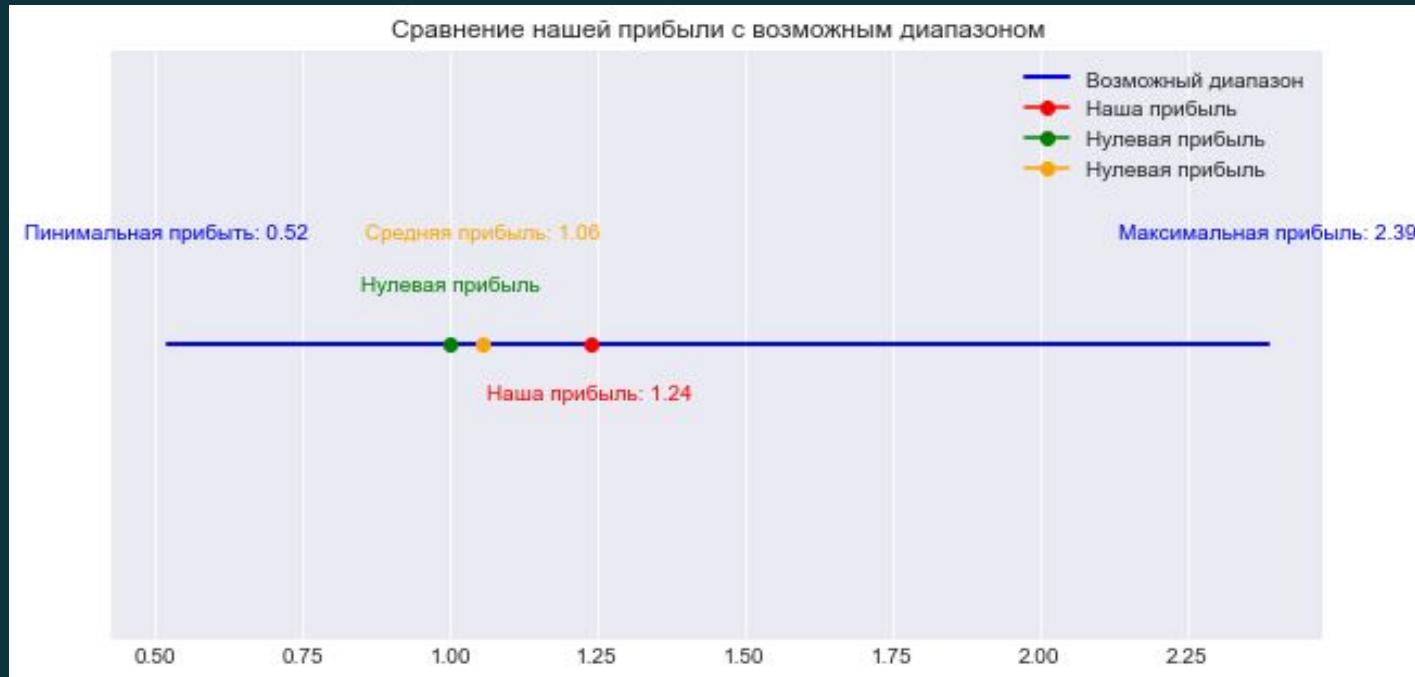
Введение:

Целью тестирования является оценка качества модели. Мы хотим узнать, насколько бы увеличился наш портфель, если бы мы использовали нашу модель для выбора активов. Для этого мы используем данные за последние 5 дней в качестве тестового набора данных, поскольку у нас ограниченное количество данных.

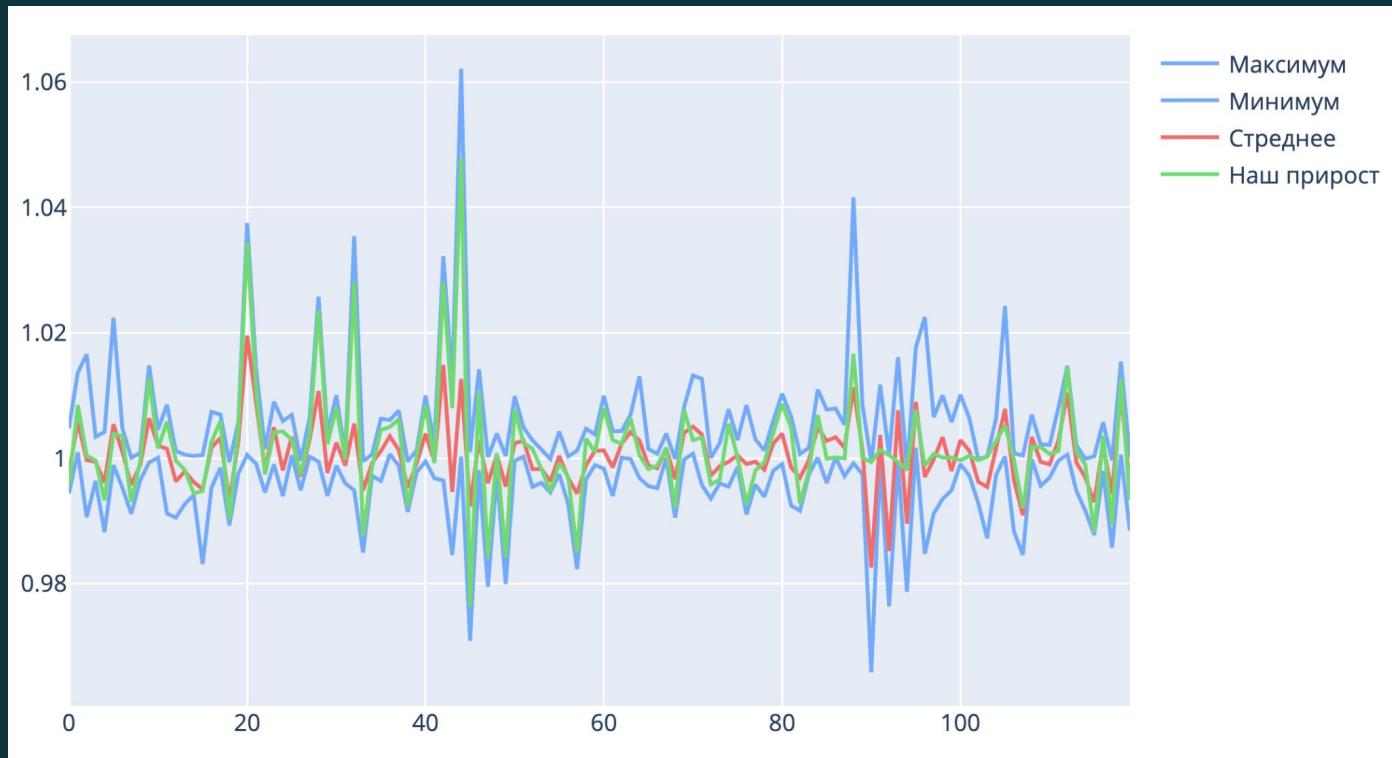
Оценка результата:

Мы сравниваем предсказания нашей модели с максимальной, минимальной и средней доходностью. Это позволяет нам оценить, насколько близко наша модель приближает оптимальный портфель к максимально возможной доходности, и насколько далеко она отстает от минимальной доходности.

Тестирование



Тестирование



**Благодайм
за внимание!**