

УДК 004.855.5

Набойченко Николай Сергеевич, магистрант, Белгородский государственный
национальный исследовательский университет, г. Белгород

Зайцева Татьяна Валентиновна, к.т.н., доцент кафедры прикладной
информатики и информационных технологий, Белгородский государственный
национальный исследовательский университет, г. Белгород

РАЗРАБОТКА МОДИФИЦИРОВАННОГО МЕТОДА ФОРМИРОВАНИЯ ОПТИМАЛЬНОГО ИНВЕСТИЦИОННОГО ПОРТФЕЛЯ С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ

Аннотация

В статье рассматривается подход к построению инвестиционного портфеля, объединяющий классическую портфельную теорию Г. Марковица с методами прогнозирования на основе градиентного бустинга. Предложенная схема включает в себя генерацию признаков из ценовых рядов, оценку ковариационной матрицы методом Леду–Вульфа и итерационное тестирование поведения портфеля на исторических данных с периодическим переобучением модели. В качестве критерия оптимальности используется коэффициент Шарпа. Тестирование на историческом периоде 2015–2026 годов показало, что портфель, составленный с помощью модифицированного метода, получил более высокую накопленную доходность и лучшее соотношение риск–доходность по сравнению с классическими подходами.

Ключевые слова: инвестиционный портфель, теория Г. Марковица, градиентный бустинг, XGBoost, оптимизация, коэффициент Шарпа.

Annotation

This article discusses an approach to constructing an investment portfolio that combines Harry Markowitz's classical portfolio theory with forecasting methods based

on gradient boosting. The proposed framework includes feature generation from price series, estimation of the covariance matrix using the Ledoit–Wolf method, and iterative backtesting of portfolio behavior on historical data with periodic model retraining. The Sharpe ratio is used as the optimality criterion. Backtesting over the historical period 2015–2026 showed that the portfolio constructed using the modified method achieved higher cumulative returns and a better risk–return ratio compared to classical approaches.

Keywords: investment portfolio, Markowitz theory, gradient boosting, XGBoost, optimization, Sharpe ratio.

Современные финансовые рынки характеризуются высокой волатильностью, нестабильностью и усложнением структуры взаимосвязей между активами. Классическая портфельная теория Гарри Марковица, являющаяся фундаментом современного инвестиционного анализа, предлагает математически строгий метод оптимизации структуры портфеля. Однако её ключевое ограничение заключается в опоре исключительно на исторические данные. Такой подход не учитывает изменение рыночных условий, паттерны поведения активов и контекстуальную информацию, что снижает предсказательную способность модели [1].

Кроме того, классическая оптимизация зачастую приводит к чрезмерной концентрации портфеля в ограниченном количестве активов, что противоречит принципу диверсификации и повышает уязвимость портфеля к рыночным шокам [3]. Указанные недостатки стимулировали разработку расширений классической модели, в том числе замену расчёта ожидаемых доходностей на основе исторических цен закрытия на прогнозные значения, полученные с помощью метода машинного обучения – градиентного бустинга.

Градиентный бустинг – метод машинного обучения, который последовательно строит ансамбль слабых прогностических моделей, как правило деревьев

решений, комбинируя их в единую сильную модель. Каждая новая модель в ансамбле исправляет ошибки предыдущих [2]. Среди преимуществ метода применительно к финансовым временным рядам – высокая устойчивость к шуму и выбросам, эффективное моделирование нелинейных зависимостей, а также возможность оценки важности признаков. Предложенный подход включает четыре этапа: сбор и предобработку рыночных данных, построение признакового описания активов, обучение прогнозной модели и решение задачи портфельной оптимизации с последующей оценкой на исторических данных.

Сбор котировок осуществляется посредством библиотеки `yfinance`, обеспечивающей доступ к скорректированным ценам закрытия. Набор инструментов охватывает десять ликвидных американских акций из различных секторов экономики; в качестве бенчмарка используется индекс S&P 500. Период наблюдений начинается с января 2015 года и продолжается до текущей даты, что обеспечивает наличие нескольких полных рыночных циклов.

Признаковое описание каждого актива формируется из трёх групп переменных. Первая группа – лаговые доходности за периоды от одного до сорока двух торговых дней. Вторая группа – скользящие оценки волатильности и импульса различных горизонтов. Третья группа – производные технического анализа: отношение текущей цены к скользящим средним, индекс относительной силы RSI с периодом 14 дней, положение цены в диапазоне 63-дневного максимума и минимума, а также скользящая оценка коэффициента бета относительно рыночного индекса. Итоговое признаковое пространство содержит порядка тридцати переменных для каждого тикера.

Центральным элементом подхода является модель прогнозирования ожидаемой доходности на горизонте h торговых дней на основе алгоритма XGBoost. Алгоритм строит ансамбль из T деревьев последовательно: каждое очередное дерево аппроксимирует отрицательный градиент функции потерь по

предсказаниям предыдущих. Итоговое предсказание – сумма вкладов всех деревьев, умноженная на шаг обучения. Для защиты от переобучения применяются ограничение глубины деревьев, случайная выборка объектов и признаков, а также L1 и L2-штрафы.

Обучение модели выполняется в скользящем режиме: на каждую дату ребалансировки модель обучается на данных за последние три торговых года. Итоговый прогноз ожидаемой доходности актива i на горизонте h формируется как взвешенная комбинация сигнала модели и исторической средней доходности:

$$\mu_i = \alpha * f_{ML}(x_i) + (1 - \alpha) * \bar{\mu}_{hist}$$

где $\alpha = 0,35$ – коэффициент смешивания; $f_{ML}(x_i)$ – предсказание модели; $\bar{\mu}_{hist}$ – выборочное среднее форвардной доходности. Полученные прогнозы ограничиваются симметричным коридором $\pm 8\%$ во избежание экстремальных значений. Для оценки ковариационной матрицы применяется сжимающий оценщик Леду–Вульфа, представляющий собой выпуклую комбинацию выборочной ковариационной матрицы и масштабированной единичной матрицы. По сравнению с выборочной матрицей этот оценщик устойчив при числе активов, сопоставимом с объёмом выборки, и значительно снижает число обусловленности матрицы. Задача портфельной оптимизации формулируется как максимизация скорректированного коэффициента Шарпа со штрафом за концентрацию:

$$\frac{w^T \mu - r_f}{\sqrt{w^T * \Sigma w}} - \lambda * w^T w \rightarrow \max$$

где w – вектор весов портфеля; μ – вектор ожидаемых доходностей; Σw – ковариационная матрица; r_f – безрисковая доходность; $\lambda = 8,0$ – коэффициент штрафа за концентрацию. Задача решается методом последовательного квадратичного программирования при ограничениях: сумма весов равна единице, минимальный вес актива – 3%, максимальный – 25%.

Для сравнения одновременно формируются два альтернативных портфеля. Классический портфель Марковица максимизирует коэффициент Шарпа без штрафа за концентрацию, используя исторические средние вместо прогнозов модели. Равновесный портфель распределяет капитал поровну между всеми активами. Тестирование проводится с минимальным начальным периодом обучения три торговых года; после его завершения модель переобучается и портфель перестраивается каждые 21 торговый день. На рисунке 1 представлена накопленная доходность сформированных портфелей, на рисунке 2 – сравнение коэффициентов Шарпа.

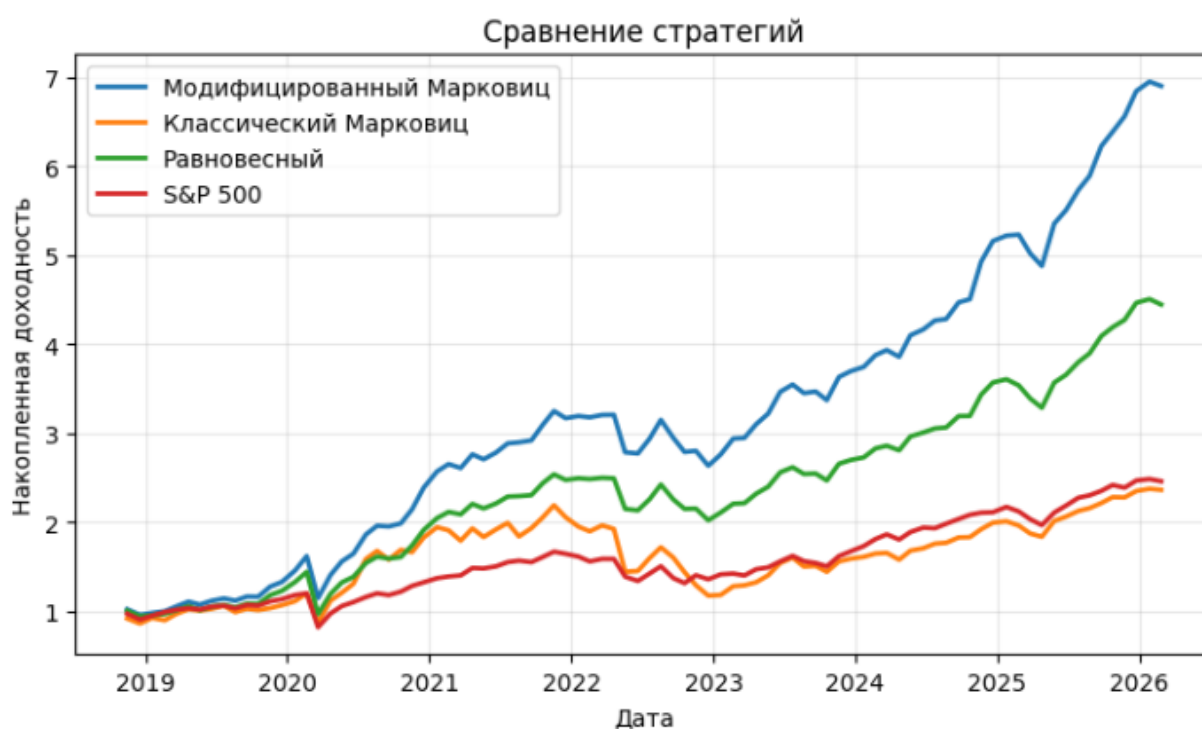


Рисунок 1 – Сравнение накопленной доходности сформированных портфелей

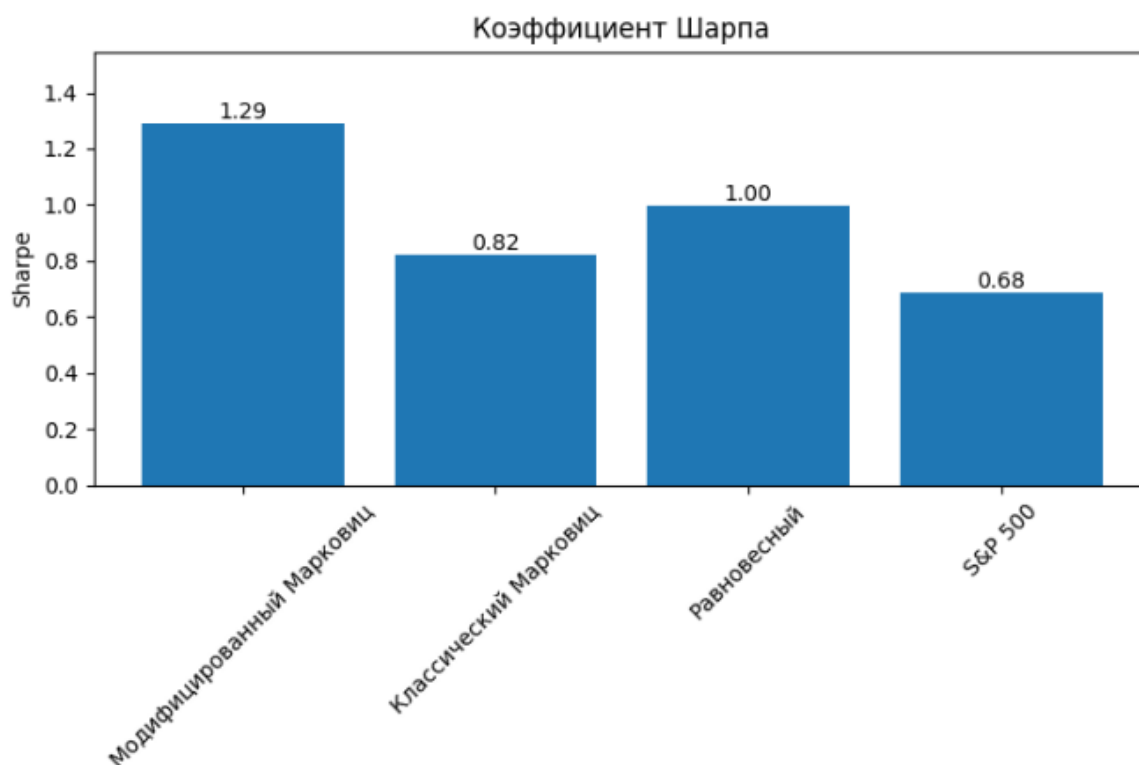


Рисунок 2 – Сравнение коэффициента Шарпа

Результаты тестирования подтверждают, что включение прогнозного сигнала модели машинного обучения в качестве источника информации об ожидаемых доходностях позволяет повысить коэффициент Шарпа по сравнению с классическими подходами. Это обосновывает целесообразность интеграции методов анализа данных в задачи портфельного управления.

Литература

1. Составление инвестиционного портфеля по Марковицу для чайников // БКС Экспресс : [сайт]. – 2018. – 7 июня. – URL: <https://bcs-express.ru/novosti-i-analitika/sostavlenie-investitsionnogo-portfelja-po-markovitsu-dlia-chainikov> (дата обращения: 06.04.2026).
2. Gradient Boosting: принципы работы и применение в машинном обучении // Sky.pro : [сайт]. – URL: <https://sky.pro/wiki/analytics/gradient-boosting-printsipy-raboty-i-primenenie-v-mashinnom-obuchenii/> (дата обращения: 08.04.2026).

3. Что такое диверсификация инвестиционного портфеля и какой портфель считается оптимальным // Т-Банк : [сайт]. – URL: <https://www.tbank.ru/invest/help/educate/how-it-works/rookie-advice/portfolio-n-risk/> (дата обращения: 12.04.2026).