

126. ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ КЛАСТЕРНОГО АНАЛИЗА АКЦИЙ БЕЛОРУССКИХ КОМПАНИЙ

Шилец О.Ю. студентка группы УИР-2

Академия управления при Президенте Республики Беларусь, г. Минск, Республика Беларусь

Шабанович Р.А. – преподаватель кафедры УИР

Аннотация. В работе показано, что иерархическая кластеризация позволяет выявить структуру белорусского рынка акций и сократить набор бумаг без потери диверсификации. По данным 13 эмитентов за 2024–2026 гг. выделено 7 кластеров и предложен портфель из 7 акций вместо 13.

Классическая теория Марковица позволяет оптимизировать соотношение риска и доходности, но не показывает, какие бумаги фактически дублируют друг друга. Для решения этой задачи применена иерархическая кластеризация, группирующая акции по сходству динамики доходностей. Цель работы – показать, как этот подход помогает описать структуру белорусского фондового рынка и сократить набор активов для портфеля на данных за 2024–2026 гг.

Проверялась гипотеза о том, что акции с высокой корреляцией доходностей относятся к одному кластеру и не должны одновременно включаться в портфель, поскольку не улучшают диверсификацию. Использованы данные ежемесячного обзора «Навигатор инвестора» [1], 16 ликвидных акций белорусских эмитентов из финансового, строительного, промышленного и энергетического секторов [2], а также реализация на Python с библиотеками pandas, numpy, scipy и matplotlib [3].

Предобработка данных и методика расчётов.

Нулевые значения заменялись предыдущими значениями с последующей интерполяцией. Для дальнейшего анализа выбраны 13 акций, имеющих не менее 10 месяцев наблюдений; СветлогорскЗЖБИИК, ЦУМ Минск и Молочный Мир исключены из-за недостатка данных. Доходности рассчитывались как логарифмические изменения цен, а годовая волатильность как стандартное отклонение доходностей, умноженное на квадратный корень из 12. Кластеризация выполнялась методом Уорда.

Первичный анализ доходности и риска.

По наблюдениям 25 месяцев рассчитаны показатели доходности и риска для 13 акций. Максимальную доходность показал Белинвестбанк (212,4 % годовых), однако для него характерен и экстремально высокий риск (283,3 %), связанный с низкой ликвидностью. Высокую доходность также продемонстрировали Трест-35 (50,5 %), СРСУ-3 (47,7 %) и БЭРН (44,3 %). Отрицательная доходность наблюдалась у Брестгазоаппарата (-18,7 %), БрестМК (-11,0 %), МозырьНПЗ (-9,3 %) и Керамина (-1,6 %).

Технологическая реализация.

Библиотеки Pandas и NumPy (Python) позволяют проводить векторизованный поиск, который выполняется на уровне низкоуровневых инструкций процессора, что в десятки раз быстрее стандартных циклов.

Создание индексов в SQL-базах данных позволяет превратить медленный линейный поиск в высокоскоростной логарифмический, что является стандартом для банковских систем.

Распределение поискового запроса по нескольким ядрам или серверам позволяет обрабатывать терабайты данных в реальном времени, что необходимо для систем обнаружения мошенничества.

Корреляционный анализ.

Корреляционный анализ выявил заметные внутрисекторные связи. Наиболее выраженные зависимости зафиксированы в строительном секторе: между Трест-35 и Минскпромстроем корреляция составила 0,44, а между Трест-35 и СРСУ-3 – 0,51 по модулю.

Результаты кластеризации.

На рисунке 1 представлены дендрограмма, корреляционная матрица, динамика цен и соотношение доходности и риска по выделенным кластерам.

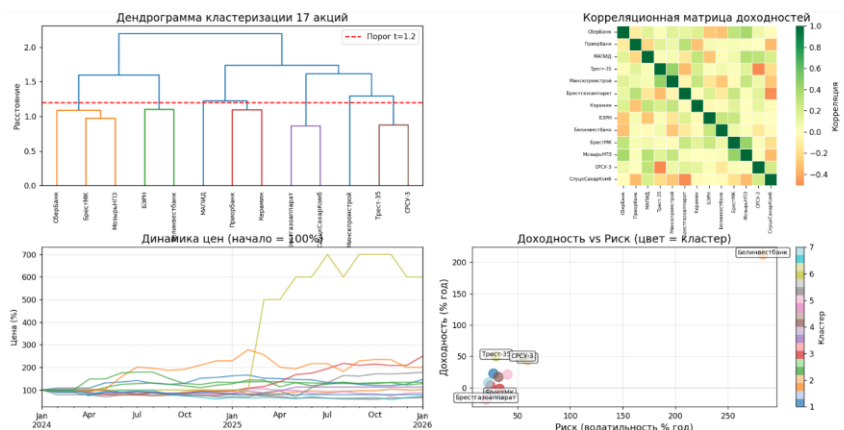


Рисунок 1 – Схема образования металл-цитратного комплекса алюминия [2], подписанная подписью выполняется с использованием

Иерархическая кластеризация позволила выделить 7 кластеров. В кластер 1 вошли СберБанк, МозырьНПЗ и БрестМК; в кластер 2 – Белинвестбанк и БЭРН; в кластер 3 – Приорбанк и Керамин; кластеры 4 и 7 образовали соответственно МАПИД и Минскпромстрой; в кластер 5 вошли СлуцкСахарКомб и Брестгазоаппарат; в кластер 6 – Трест-35 и СРСУ-3. Полученная структура в целом соответствует отраслевым связям между эмитентами и выделяет группы с близким профилем риска и доходности.

Практический вывод состоит в выявлении избыточных бумаг внутри кластеров. Например, в паре Трест-35 - СРСУ-3 предпочтителен Трест-35, так как при сопоставимой доходности он имеет меньший риск. Кластеры, представленные одиночными эмитентами, напротив, отражают уникальное поведение бумаг и важны для диверсификации.

На основе кластерного анализа сформирован портфель, включающий по одной акции из каждого кластера. В итоговый набор вошли:

1. СберБанк (доходность 22,6 %, риск 27,0 %);
2. Белинвестбанк (212,4 %; 283,3 %);
3. Приорбанк (3,5 %; 24,0 %);
4. МАПИД (17,2 %; 31,8 %);
5. СлуцкСахарКомб (21,1 %; 40,4 %);
6. Трест-35 (50,5 %; 29,6 %);
7. Минскпромстрой (8,1 %; 22,1 %).

Это позволяет сократить исходный набор из 13 бумаг до 7, то есть на 46 %, без потери диверсификационных свойств и делает кластеризацию удобным этапом перед последующей оптимизацией по Марковицу.

С точки зрения развития рынка выявленная структура показывает его отраслевую неравномерность. Наиболее отчетливо выделяются строительный и финансовый сегменты, тогда как ряд кластеров представлен одиночными эмитентами. Это указывает на ограниченные возможности диверсификации и подчёркивает необходимость расширения круга ликвидных бумаг из разных отраслей.

Для инвестора такой результат означает, что диверсификацию следует строить по кластерам, а не по числу бумаг. Для регулятора такая структура показывает ограничения рынка, повышающие его уязвимость к локальным шокам.

Выводы.

Иерархическая кластеризация показала пригодность для предварительного отбора активов на белорусском фондовом рынке. Выделенные 7 кластеров отражают взаимосвязи между 13 эмитентами, уменьшают размерность задачи и помогают формировать более содержательно диверсифицированный портфель. Перспективы дальнейшей работы связаны с включением облигаций, макроэкономических факторов и использованием более сложных алгоритмов кластеризации.

Список использованных источников:

1. Навигатор инвестора / Aigenis. / 2024–2026. – URL: aigenis.by/navigator-investora/ (дата обращения: 20.03.2026).
2. Васильев, В. А. Математические модели оценки и управления финансовыми рисками хозяйствующих субъектов / В. А. Васильев, А. В. Летчиков, В. Е. Лялин // Экономика и управление: проблемы, решения. – 2023. – Т. 1, № 5. – С. 219–225.
3. Pedregosa F. Scikit-learn: Machine Learning in Python / F. Pedregosa et al. // Journal of Machine Learning Research. – 2011. – Vol. 12. – P. 2825–2830.