

ИСПОЛЬЗОВАНИЕ КЛАСТЕРНОГО АНАЛИЗА ДЛЯ ФОРМИРОВАНИЯ ПОРТФЕЛЯ ЦЕННЫХ БУМАГ В ИНВЕСТИЦИОННЫХ СИСТЕМАХ (РОБОЭДВАЙЗЕРАХ)

Тюхова Е.М.,
Katholieke Universiteit Leuven
lena.tyukhova@gmail.com

Сизых Д.С.
Институт проблем управления им. В.А. Трапезникова РАН
Институт Национальный исследовательский университет «Высшая школа экономики» (НИУ ВШЭ)
D.Sizykh@gmail.com

Аннотация: В работе проанализирована возможность использования кластерного анализа для формирования эффективного портфеля ценных бумаг. Выявлены особенности использования кластерного анализа и апробирован алгоритм его применения на примере формирования портфеля акций робоэдвайзером.

Ключевые слова: кластерный анализ, котировки акций, формирование портфеля ценных бумаг, метод Марковица, динамические показатели, рыночные характеристики акций.

Введение

Расширение инвестиционной деятельности способствует развитию и внедрению различных автоматизированных систем с целью повышения ее эффективности. В последние годы значительно возрастают предложения по внедрению и использованию разнообразных робоэдвайзеров для формирования инвестиционных портфелей ценных бумаг, для инвестиционного консультирования и пр. Современные робоэдвайзеры применяются как конечными пользователями-инвесторами, так и менеджерами инвестиционных управляющих компаний, как профессиональными инвесторами, так и непрофессиональными. В связи с этим растет потребность в постоянном совершенствовании математических методов и моделей, используемых в робоэдвайзерах для предварительной обработки и анализа различной информации с целью повышения качества предоставляемых услуг по инвестиционным консультациям, по формированию инвестиционных портфелей, выбору ценных бумаг, перебалансировке и оптимизации портфелей и пр. [1, 3-6]

Самым востребованным направлением консультаций с помощью робоэдвайзеров является формирование и перебалансировка портфелей ценных бумаг. В данном направлении уже разработаны различные алгоритмы выбора ценных бумаг. Некоторые алгоритмы нацелены только на выбор массива ценных бумаг и предоставление возможности инвесторам самим выбрать те ценные бумаги, которые они хотели бы включить в формируемый портфель. Есть алгоритмы, которые предлагают инвесторам варианты с уже сформированными портфелями.

Актуальным в настоящее время является разработка и внедрение различных интеллектуальных алгоритмов, которые способствуют формированию эффективных и оптимальных инвестиционных

портфелей. Одним из таких инновационных направлений является анализ возможности использования кластерного анализа. Известно, что для формирования эффективного инвестиционного портфеля необходимо включение в него ценных бумаг, которые имеют противоположные стоимостные тенденции. В данном случае использование кластерного анализа является целесообразным, поскольку он направлен на разбиение объектов по совокупности их признаков на однородные группы. Цель кластерного анализа, как метода изучения однородности сложных и неочевидно взаимосвязанных объектов, состоит в выделении кластеров из исследуемой совокупности объектов. Поэтому выбор ценных бумаг из разных кластеров и будет способствовать снижению рисков инвестиционного портфеля [2,4,7].

В работе исследована возможность использования кластерного анализа для предварительного формирования групп ценных бумаг, из которых в дальнейшем формируется портфель, разработан и предложен алгоритм. Такой предварительный этап предоставляет инвесторам сделать достаточно эффективный выбор уже на начальном этапе. В дальнейшем могут использоваться различные алгоритмы оптимизации.

1 Анализ методов формирования портфелей ценных бумаг

Первым, кто постарался формализовать выбор инвестора, был Гарри Марковиц, который в 1952 году сформулировал модель формирования оптимального инвестиционного портфеля: инвестору необходимо диверсифицировать инвестиционный портфель таким образом, чтобы максимизировать ожидаемую доходность данного портфеля при заданном уровне риска, или же минимизировать риск при заданной доходности. Преимуществом данного подхода является то, что он позволяет добиться оптимального соотношения доходности и риска, а также построить портфель на принципе стабильности: портфель ребалансируется в соответствии с доходностью входящих в него активов во избежание игры на колебаниях. Недостаток данной теории заключается в том, что сила связи между ценными бумагами измеряется только с помощью коэффициента корреляции. Данный показатель очень чувствителен к колебаниям на финансовом рынке, что может привести к тому, что сформированный на его основе портфель ценных бумаг будет нестабилен в периоды высокого риска [4].

Возможность построения кластеров ценных бумаг была предложена Ж. Рен. В своей работе [7] он предлагает использовать метод формирования кластеров ценных бумаг на основании силы связи между ними. Сила связи здесь так же, как и в методе Марковица, измеряется с помощью корреляции. В работе предлагается формирование кластеров из ценных бумаг, корреляция между которыми выше определенного порогового значения (0.2). В дальнейшем данные кластеры используются как самостоятельные «ценные бумаги», к которым применяются оптимизационные методы Марковица для составления оптимального инвестиционного портфеля. Так как ценные бумаги с высокой корреляцией объединяются в отдельные кластеры, то корреляция между самими кластерами минимальна, что позволяет формировать менее рискованные инвестиционные портфели. Достоинством данной работы является то, что здесь впервые предлагается использовать подход с формированием кластеров. Более того, использование корреляции здесь оправдано тем, что намного сложнее оценить будущую доходность ценных бумаг, чем оценить корреляцию между ними. Тем не менее данная работа не лишена недостатков: как уже было сказано выше, корреляция сильно меняется в периоды финансовой нестабильности, а значит сформированные в соответствии с ней кластеры окажутся неадекватно структурированными в некоторые периоды времени.

Подход с использованием кластеризации для портфельного анализа также рассмотрен К. Марвин. В своей работе [5] она предлагает использование непосредственно кластерного анализа для формирования кластеров, в которые попадают ценные бумаги тех компаний, для которых показатели Revenues/Assets и Net Income/Assets наиболее близки. Далее из каждого кластера выбирается ценная бумага по наибольшему коэффициенту Шарпа и включается в итоговый инвестиционный портфель. Достоинством данного подхода является то, что здесь кластеры формируются по показателям финансовой деятельности компаний, что позволяет более точно определить степень похожести/непохожести ценных бумаг. Недостатком данной работы является тот факт, что для измерения связи между ценными бумагами используется всего два показателя финансовой деятельности, которые могут недостаточно полно отражать общую ситуацию в компании, а поэтому сформированные кластеры могут не учесть некоторых важных сходств или различий.

Рынок ценных бумаг и инвестиций динамично развивается в связи с развитием информационных технологий и искусственного интеллекта. Финансовые консультанты на рынке

все чаще заменяются автоматизированными системами (робот-эдвайзерами), способными предоставлять финансовые консультации, формировать портфель ценных бумаг и управлять им без участия человека (или при минимальном его участии). Внедрение автоматизированных систем формирования инвестиционного портфеля на рынке ценных бумаг требует большой осторожности, а также учета недостатков таких систем. В частности, необходимо сформировать развитую финансовую инфраструктуру, а также повсеместно повышать финансовую грамотность инвесторов и их готовность использовать робот-эдвайзеры. В работе [3] отмечается, что в процессе развития рынка робот-эдвайзеров возникает множество трудностей и недостатков, которые необходимо принимать во внимание при их внедрении в информационную инфраструктуру компании.

2 Особенности применения кластерного анализа для формирования портфеля ценных бумаг

С помощью кластерного анализа формируем кластеры со сходными объектами (компаниями, акции которых рассматриваются для включения в портфель) [6,7]. Для оценки меры схожести объектов используем следующие признаки:

- котировки стоимости акций (рассматриваются временные ряды);
- различные динамические показатели изменения котировок акций;
- основные рыночные мультипликаторы деятельности компаний-эмитентов акций.

Таким образом, исходные данные при проведении кластерного анализа задаются матрицей признаков, в которой каждому объекту соответствует определенный набор признаков [1,2,4,7]

$$X = \{x_{ij}\}, (i = 1, 2, \dots, n), (j = 1, 2, \dots, m)$$

где x_{ij} – j -ый признак i -того объекта;

m – количество признаков, характеризующих исследуемые объекты;

n – количество исследуемых объектов для кластеризации.

В качестве используемой метрики для оценки сходства объектов используется расстояние между объектами, т.е. $d(x, y)$ – расстояние между объектами x и y в пространстве признаков, характеризующих объекты кластеризации с введенной метрикой. В качестве метрики используется квадрат евклидова расстояния:

$$d(x, y) = \sum_{j=1}^m (x_j - y_j)^2,$$

где m – количество признаков $j=1, 2, \dots, m$

Используется метод кластеризации k -средних. Количество кластеров определяется по количеству обобщенных факторов, выделенных с помощью упрощенного варианта факторного анализа, который проводится методом главных компонент с собственными значениями более 0,8. Поскольку с помощью данного метода выделяются обобщенные факторы (признаки группируются по показателю взаимной корреляции), то можно предположить, что количество кластеров может соответствовать количеству выделенных факторов. В кластерном анализе объекты группируются по расстоянию между ними, а мерой расстояния выступают значимые признаки. Предполагаем, что в выделенных обобщенных факторах найдется, по крайней мере, один значимый признак, который будет участвовать в формировании кластеров. Взаимное сочетание таких признаков, как правило, будет повышать количество кластерных группировок.

В качестве оценки связанности кластеров используют отношение среднего внутрикластерного расстояния к межкластерному:

$$Q_{i,j} = \frac{Q1_i + Q1_j}{2Q2_{i,j}}$$

где $Q1_i$ и $Q1_j$ – средние внутрикластерные расстояния классов i и j ;

$Q2_{i,j}$ – среднее межкластерное расстояние между этими же кластерами.

Для оценки значимости различий применяется t -критерий Стьюдента.

Формально задача кластерного анализа представляет собой минимизацию среднеквадратического отклонения для объектов каждого кластера:

$$\arg \min \sum_{i=1}^m \sum_{x \in S_i} d(P, \mu_i)^2,$$

где μ_i – начальные центры кластеров,

$$i = 1, \dots, m,$$

$d(P, \mu_i)$ – функция расстояния между объектом P и центром кластера μ_i .

Алгоритм кластерного анализа:

- Инициализация начальных центров кластеров произвольным образом: произвольным образом выбирается множество начальных центров кластеров $\mu_i, i = 1, \dots, m$.
- Распределение объектов по кластерам: на шаге $t \forall P_i \in P, i = 1, \dots, n : P_i \in g_j \Leftrightarrow j = \arg \min d(P_i, \mu_m^{(t-1)})^2$.
- Пересчет центров кластеров после отнесения объектов к кластерам: на шаге t также корректируются центры кластеров с учетом добавленных в кластеры объектов: $\forall i = 1, \dots, m : \mu_i^{(t)} = \frac{1}{|g_i|} \cdot \sum_{P \in g_i} P$. Благодаря шагам распределения объектов по кластерам и пересчета центров кластеров уменьшается сумма квадратов расстояний.
- Проверка условия продолжения работы алгоритма кластеризации: если после пересчета центра кластеров на шаге $t \mu_i^{(t)} \neq \mu_i^{(t-1)}, i = 1, \dots, m$, то есть после пересчета центры кластеров изменились, то работа алгоритма продолжается. Если $\mu_i^{(t)} = \mu_i^{(t-1)}, i = 1, \dots, m$, то есть после пересчета центры кластеров остались прежними, то работа алгоритма прекращается [1,2].

Множеством объектов в кластерном анализе для портфельного анализа является множество ценных бумаг компаний из отраслей информационных технологий и телекоммуникаций. В качестве множества наблюдаемых характеристик объектов рассматриваются три разных множества: множество, состоящее из месячных котировок ценных бумаг; множество динамических показателей ценных бумаг; множество, состоящее из рыночных показателей и финансовых мультипликаторов.

Известно, что чем больше различаются объекты инвестиций между собой, тем ниже показатели риска, при включении этих объектов в один портфель. При этом данные объекты могут иметь различную доходность. Поэтому, после получения набора кластеров появляется задача формирования правила выбора объектов инвестирования для формирования портфеля ценных бумаг. Формирование портфеля на базе ценных бумаг из разных кластеров будет снижать возможные риски. Что касается доходности данного объекта, то она может быть разной не зависимо от положения объекта в кластере.

Поэтому формирование портфеля ценных бумаг с помощью данных кластерного анализа будем проводить по следующему правилу: выбираем ценные бумаги из разных кластеров с учетом максимально доходности при вероятности принадлежности объекта к данному кластеру более 0,7. При использовании кластерного анализа в робоэдвайзерах инвестору предъявляются выбранные ценные бумаги по кластерам с указанием доходности и вероятности их принадлежности данному кластеру, и инвестор сам может указать те из них, которые он хотел бы иметь в своем портфеле.

В случае формирования портфеля робоэдвайзером, автоматически выбирается от 2 до 10-ти ценных бумаг из разных кластеров с учетом максимально доходности и вероятности принадлежности объекта к данному кластеру более 0,7.

После выбора ценных бумаг для формирования портфеля, проводится процесс оптимизации. В качестве модели оптимизации в данном случае рекомендуется использовать метод Марковица для построения инвестиционного портфеля максимальной доходности: [4]

$$\left\{ \begin{array}{l} \sum_{i=1}^n w_i \cdot r_i \rightarrow \max \\ \sqrt{\sum_{i=1}^n \sum_{j=1}^n w_i w_j \text{Cov}_{ij}} < \sigma_p \\ \sum_{i=1}^n w_i = 1 \\ w_i \geq 0 \end{array} \right.$$

Доходность портфеля Марковица рассчитывается как средневзвешенная доходность ценных бумаг, входящих в портфель:

$$r_p = \sum_{i=1}^n w_i \cdot r_i,$$

где w_i – удельный вес ценной бумаги в портфеле,

r_i - доходность i -ой ценной бумаги за определенный период (в данном случае, за год).

Риск в портфеле Марковица рассчитывается через риск входящих в него активов и взаимной зависимости доходностей этих активов:

$$\sqrt{\sum_{i=1}^n \sum_{j=1}^n w_i w_j Cov_{ij}},$$

где w_i – удельный вес i -ого актива в инвестиционном портфеле,

w_j - удельный вес j -ого актива в инвестиционном портфеле,

Cov_{ij} – ковариация доходностей i -ого и j -ого актива.

При формировании инвестиционного портфеля Марковица на веса ценных бумаг накладывается ограничение сверху: вес не более заданного значения для обеспечения диверсификации портфеля и отсутствия сильного доминирования ценных бумаг только одной компании в портфеле.

3 Практические результаты исследования

Практическая апробация применения кластерного анализа для формирования инвестиционных портфелей ценных бумаг проводилась на примере формирования отраслевого портфеля акций. Проведено формирование портфелей акций интернет компаний и телекоммуникационных компаний. Рассматривались акции 30 интернет компаний и 45 телекоммуникационных компаний. Исследовались данные ежемесячно в течение года за период 2015 – 2018 года.

Исследовалась эффективность формирования портфелей по данным кластерного анализа, который проводился по следующим трем группам признаков: по котировкам акций (как временным рядам стоимости акций); по динамическим показателям изменений стоимости акций в течение года; по рыночным мультипликаторам компаний-эмитентов акций (данные взяты из базы Блумберг). В качестве метода кластерного анализа использовался k -средних. Количество кластеров определялось по результатам факторного анализа по количеству обобщенных факторов. При использовании в качестве признаков котировок акций, факторный анализ определял факторы по характеру изменений котировок акций за год. Основные показатели, которые были отобраны как динамические характеристики котировок акций:

- амплитуда (размах) колебаний котировок;
- значение 1% прироста котировок (отношение абсолютного прироста к темпу прироста);
- среднее квадратическое отклонение котировок;
- линейный коэффициент вариации (доля усредненного значения абсолютного отклонения котировок акций от средней величины);
- колеблемость по абсолютной величине (обобщенная характеристика степени колеблемости значений котировок в течение года);
- коэффициент вариации, коэффициент колеблемости (однородность совокупности показателей определяется как отношение среднего квадратического отклонения к среднему уровню временного ряда за весь период);
- коэффициент устойчивости показывает обеспечение устойчивости уровней ряда относительно тренда (определяется в % как 100 - коэффициент колеблемости: чем ближе к 0, тем менее устойчивый процесс изменений котировок акций);
- индекс устойчивости (отношение средних уровней котировок акций за периоды роста к средним уровням за периоды снижения: чем ближе значение индекса к единице, тем выше устойчивость);
- коэффициент осцилляции (относительная колеблемость крайних значений котировок акций вокруг среднего их значения);
- коэффициент Спирмена (устойчивость тенденции изменений котировок акций как временного ряда);
- коэффициент стабильного роста котировок акций (стабильность роста котировок акций);
- индекс относительной силы, Relative Strength Index RSI (измеряет скорость и величину движения направленной цены и отображает данные, колеблясь между 0 и 100);
- показатель Hurst (случайность характера изменений котировок акций).

В качестве рыночных мультипликаторов использовались следующие показатели: P / Tangible Book; P / E; P / Book; P / Sales; P/Cash Flow; P/Free Cash Flow; PEG; DFL; EPS; доходность акций; риск доходности; Enterprise Value; Market Capitalization.

По данным кластерного анализа формировался инвестиционный портфель по следующему алгоритму: выбираем ценные бумаги из разных кластеров с учетом максимально доходности при вероятности принадлежности объекта к данному кластеру более 0,7. При формировании портфеля учитывались ограничения по весам: не более 10% акции одного вида. На основе построенного портфеля была проведена оптимизация по модели Марковица с ограничением по весам: не более 20% акции одного вида. Показатели эффективности портфелей оценивались по показателям доходности, риска, коэффициентам Бета, Трейнора, Йенсена и Модильяни.

Таким образом, было сформировано по два инвестиционных портфеля на основе результатов кластерного анализа по выбранным трем признакам. Один портфель формировался без оптимизации, а другой с оптимизацией по методу Марковица. При этом формирование портфелей проводилось по данным за год из рассматриваемых периодов (2015г., 2016г., 2017г.) отдельно для акций интернет компаний и акций телекоммуникационных компаний.

Проведем сравнение эффективности полученных портфелей. Определим наиболее оптимальный вариант проведения кластерного анализа относительно используемых признаков.

Таблица 1. Структура инвестиционного портфеля, сформированного по данным кластерного анализа по котировкам акций за 2017 год

Компания	Кластер	Расстояние до центра кластера, ед.	Годовая доходность, %	Доля акций в портфеле по Марковицу, ед.
EDreams ODIGEO	4	0	47,2	0,2
Naver Corporation	1	0	11,6	0
salesforce.com Inc (CRM US)	3	18,4	40,1	0,2
Cimpres N.V.	3	41,6	26,9	0
ASOS.com	3	72,7	24,3	0
Fiserv Inc	3	87,5	21,0	0
Amadeus IT Group	3	119,1	33,1	0,2
Amazon.com Inc	2	899,0	44,4	0,2
Alphabet Inc	2	1066,9	30,4	0,2
Booking Holdings	2	2027,7	17,0	0

Согласно таблице 1 в портфель без оптимизации включены акции десяти компаний, доля каждой 10%. В портфель с учетом оптимизации по Марковицу включены акции пяти компаний доля каждой 20%.

Сравним инвестиционные портфели, сформированные с применением метода кластеризации.

Таблица 2. Показатели эффективности инвестиционных портфелей акций интернет-компаний, сформированных по результатам кластерного анализа с проведением оптимизации за период 2015-2017 г.г.

Признаки акций, используемые для кластеризации	Доходность портфеля, %	Риск портфеля, %	Коэффициент Бета, ед.	Коэффициент Шарпа, ед.	Коэффициент Трейнора, ед.	Коэффициент Альфы Йенсена, ед.	Коэффициент Модильяни, ед.
Динамические показатели, 2015	41,8	18,1	1,162	1,801	0,281	0,369	0,172
Котировки, 2015	49,1	21,0	1,145	1,900	0,348	0,441	0,176

Признаки акций, используемые для кластеризации	Доходность портфеля, %	Риск портфеля, %	Коэффициент Бета, ед.	Коэффициент Шарпа, ед.	Коэффициент Трейнора, ед.	Коэффициент Альфы Йенсена, ед.	Коэффициент Модильяни, ед.
Мультипликаторы, 2015	48,4	18,0	0,912	2,184	0,430	0,426	0,189
Динамические показатели, 2016	26,5	9,6	1,093	2,587	0,228	0,187	0,123
Котировки, 2016	22,9	11,5	0,970	1,843	0,219	0,158	0,092
Мультипликаторы, 2016	16,3	3,9	0,834	3,799	0,176	0,100	0,173
Динамические показатели, 2017	42,7	5,6	1,369	9,597	0,393	0,045	0,033
Котировки, 2017	39,1	6,4	1,370	7,830	0,367	0,008	0,006
Мультипликаторы, 2017	52,7	24,5	0,992	2,604	0,644	0,281	-0,073

Анализ показателей эффективности показывает, что портфели, сформированные с применением кластерного анализа и последующей оптимизации Марковица, являются оптимальными как по доходности (средняя доходность составляет ~38%), так и по риску (средний риск составляет ~12%). В соответствии с коэффициентом Бета стратегия инвестора консервативна, а риск низкий для портфелей, сформированных по данным финансовых мультипликаторов; кластеризация по динамическим показателям и котировкам акций способствует формированию высокорискованных портфелей ценных бумаг при агрессивной стратегии управления. Коэффициент Шарпа для всех сформированных портфелей выше единицы, что говорит об их высокой эффективности. Коэффициенты Трейнора и Альфы Йенсена положительны для всех портфелей, а коэффициент Модильяни сравнительно высок [1].

Таблица 3. Показатели эффективности инвестиционных портфелей акций интернет-компаний, сформированных только по результатам кластерного анализа без проведения оптимизации за период 2015-2017 г.г.

Признаки акций, используемые для кластеризации	Доходность портфеля, %	Риск портфеля, %	Коэффициент Бета, ед.	Коэффициент Шарпа, ед.	Коэффициент Трейнора, ед.	Коэффициент Альфы Йенсена, ед.	Коэффициент Модильяни, ед.
Динамические показатели, 2015	28,2	19,3	1,121	0,981	0,169	0,231	0,136
Котировки, 2015	31,8	23,3	1,112	0,970	0,203	0,267	0,135
Мультипликаторы, 2015	34,6	19,3	0,990	1,317	0,256	0,290	0,151
Динамические показатели, 2016	17,7	11,2	1,100	1,450	0,147	0,099	0,076

Признаки акций, используемые для кластеризации	Доходность портфеля, %	Риск портфеля, %	Коэффициент Бета, ед.	Коэффициент Шарпа, ед.	Коэффициент Трейнора ед.	Коэффициент Альфы Йенсена, ед.	Коэффициент Модильяни, ед.
Котировки, 2016	15,2	11,5	1,002	1,189	0,136	0,080	0,065
Мультипликаторы, 2016	11,4	5,9	0,913	1,662	0,108	0,047	0,084
Динамические показатели, 2017	32,4	12,1	0,981	3,605	0,444	0,082	-0,058
Котировки, 2017	29,6	11,2	0,950	3,654	0,429	0,065	-0,057
Мультипликаторы, 2017	35,6	22,5	0,769	2,080	0,608	0,191	-0,081

Анализ показателей эффективности показывает, что портфели, сформированные с применением кластерного анализа, являются эффективными по показателям доходности (средняя доходность составляет ~26%) и риска (средний риск составляет ~15%). Практически для всех сформированных портфелей стратегия инвестора агрессивна в соответствии с коэффициентом Бета. Коэффициент Шарпа выше единицы для всех портфелей, за исключением портфелей по динамическим показателям и котировкам для 2015 года, что говорит, что сформированные портфели эффективны, но в некоторых случаях их доходность не оправдана слишком высоким риском. Коэффициенты Трейнора и Альфы Йенсена положительны, что говорит об эффективной стратегии инвестора. Коэффициент Модильяни положителен для всех портфелей, за исключением 2017 года, что свидетельствует об эффективном портфельном управлении [1].

Сравнить полученные портфели ценных бумаг помогает визуализация их показателей эффективности.

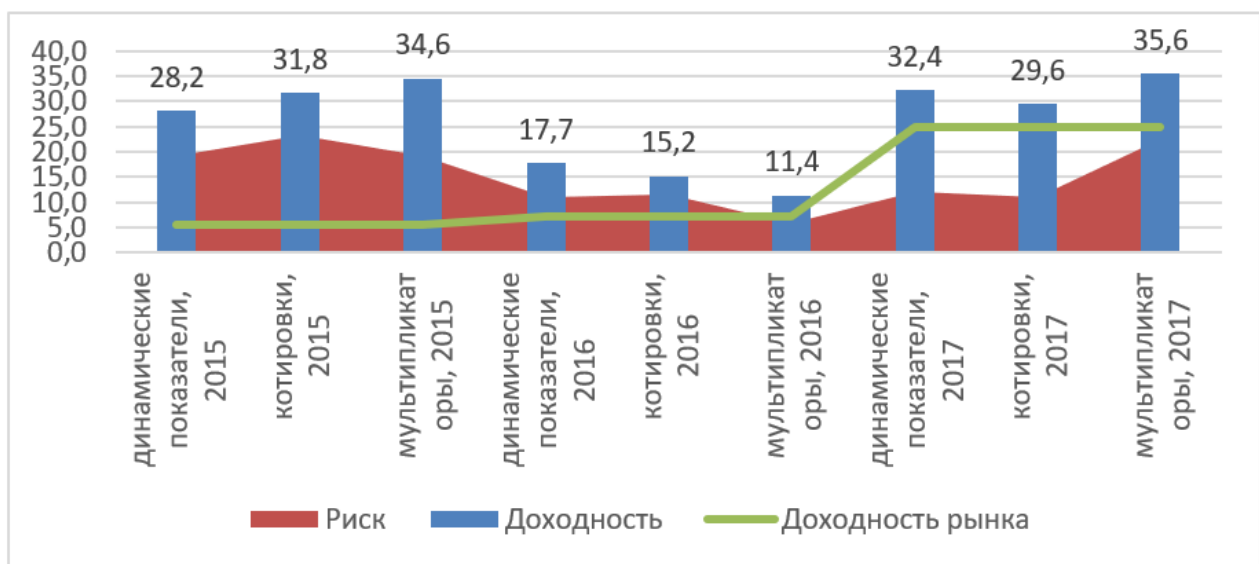


Рис. 1. Диаграмма показателей риска и доходности инвестиционных портфелей акций интернет-компаний, сформированных только по результатам кластерного анализа без проведения оптимизации за период 2015-2017 г.г.

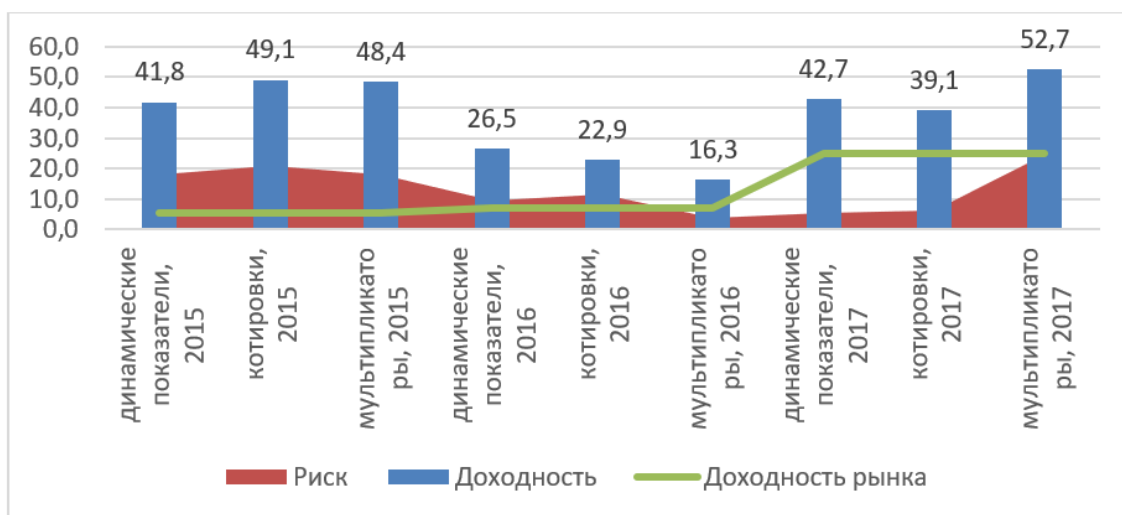


Рис. 2. Диаграмма показателей риска и доходности инвестиционных портфелей акций интернет-компаний, сформированных по результатам кластерного анализа с проведением оптимизации за период 2015-2017 г.г.

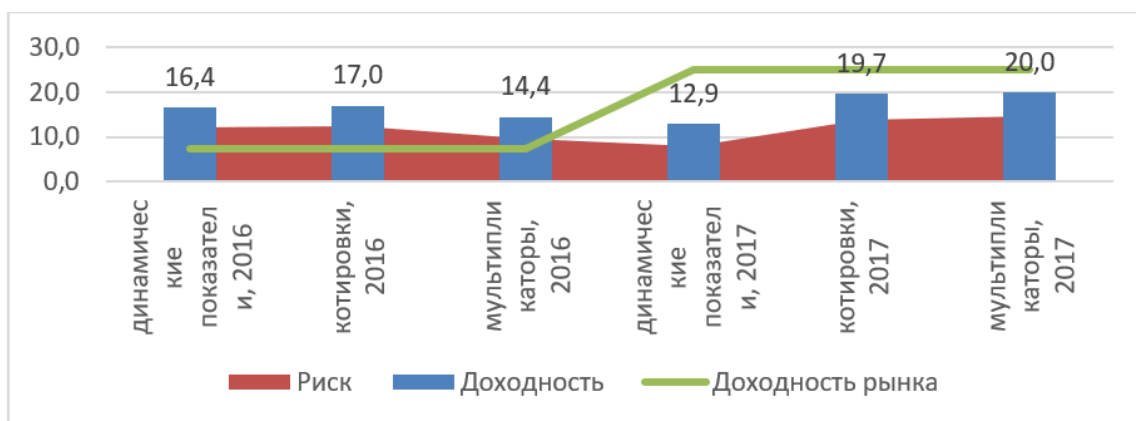


Рис. 3. Диаграмма показателей риска и доходности инвестиционных портфелей акций телекоммуникационных компаний, сформированных только по результатам кластерного анализа без проведения оптимизации за период 2016-2017 г.г.

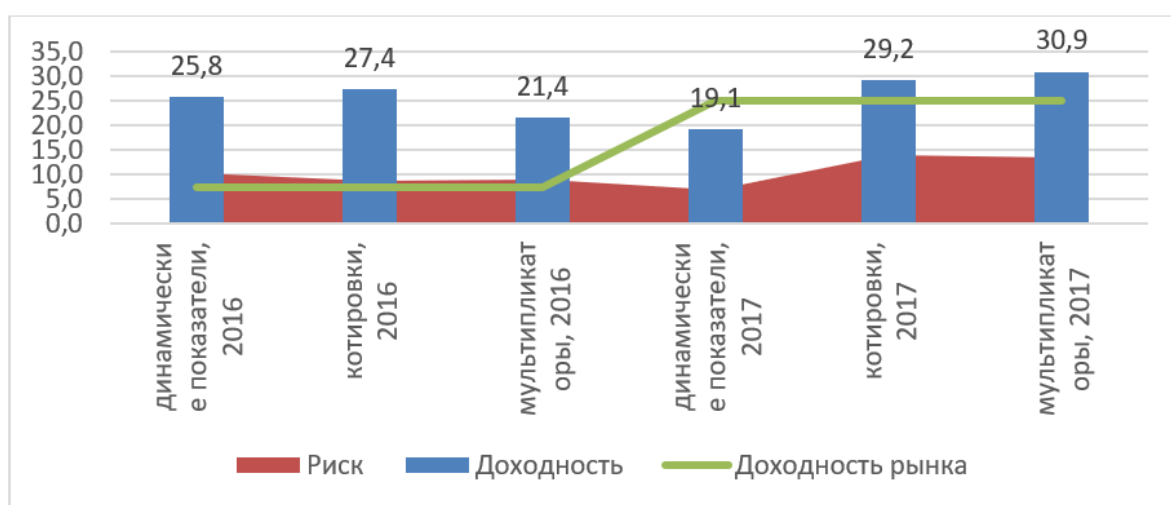


Рис. 4. Диаграмма показателей риска и доходности инвестиционных портфелей акций телекоммуникационных компаний, сформированных по результатам кластерного анализа с проведением оптимизации за период 2016-2017 г.г.

Сравнительный анализ инвестиционных портфелей акций сформированных на базе данных кластерного анализа показал, что:

1. использование данных кластерного анализа позволяет формировать эффективные инвестиционные портфели, повысить их доходность;
2. данные кластерного анализа можно использовать как предварительный этап для формирования эффективных портфелей;
3. доходность портфелей, сформированных на основе данных кластерного анализа с дальнейшей оптимизацией, выше доходности рынка. Даже если портфель сформирован только по данным кластерного анализа без оптимизации, то его доходность, как правило выше рынка;
4. использование оптимизации Марковица для портфеля, сформированного по данным кластерного анализа, позволяет повысить доходность и снизить риск: доходность повышается, как правило, от 30% и выше. При этом риск портфеля снижается, а доходность повышается.

По проведенным исследованиям однозначного ответа относительно выбора признаков, по которым наилучшим образом разделяются акции по группам для дальнейшего формирования инвестиционного портфеля, не получено. При этом в большинстве случаев наиболее эффективный портфель формируется по данным кластеризации акций по котировкам.

Заключение

Проведена оценка возможности применения кластерного анализа для процесса формирования портфеля ценных бумаг. Анализ показал, что кластеризация позволяет сформировать более эффективные, чем рыночные индекс, портфели ценных бумаг. При этом последующая оптимизация Марковица на основе данных кластерного анализа позволяет повысить доходность портфеля в среднем на 10-11% и снизить риск на 2-3%. В целом, использование кластерного анализа с последующей оптимизацией Марковица позволяет повысить доходность в среднем на 25% над рыночной.

Более эффективно применять кластерный анализ как предварительный этап с последующей оптимизацией портфеля по методу Марковица. Процесс проведения кластеризации акций может проводиться по предложенному в данной работе алгоритму, а в качестве используемых признаков рекомендуется применение котировок акций.

С большой долей вероятности можно утверждать, что не только акции, но и любые ценные бумаги с целью формирования эффективного инвестиционного портфеля могут быть разделены по кластерам аналогичным образом.

Полученный алгоритм применения кластерного анализа для формирования инвестиционных портфелей является достаточно простым и может применяться в робоэдвайзерах с целью интеллектуализации их работы. При выполнении данного исследования был составлен алгоритм и написан экспериментальный вариант программы для бота, позволяющий формировать портфель акций с применением кластерного анализа. Разработаны рекомендации для автоматизации процесса формирования портфеля ценных бумаг с использованием данных кластерного анализа: разработан алгоритм автоматизации, визуализированный с помощью блок-схем, который может быть использован при разработке автоматизированной системы практически на любой платформе. В качестве примера автоматизации разработан робо-эдвайзер на базе Telegram-бота, логика которого спроектирована на основе разработанного ранее алгоритма с применением данных кластерного анализа.

Особенностью данного исследования является тот факт, что оно несет прикладной характер и ориентировано на удовлетворение растущих потребностей рынка в автоматизации инвестиционных процессов. Исследование имеет большое значение как для инвесторов различной квалификации, так и для крупных инвестиционных компаний. Робо-консультант – отличное решение для тех, кто хочет автоматически формировать свой портфель и управлять им. Разработанная система может быть внедрена в IT-инфраструктуру инвестиционных компаний как дополнительный источник получения прибыли и привлечения новых клиентов, а разработанные рекомендации по автоматизации могут лечь в основу автоматизированных систем на разных платформах, что позволит упростить инвестиционный процесс и сделать его более эффективным.

Кроме того, исследование будет полезно для дальнейшего развития теоретических работ, посвященных использованию кластерного анализа для формирования оптимального

инвестиционного портфеля. Автоматизирован может быть также процесс ребалансировки портфеля ценных бумаг, что позволит иметь оптимальный портфель в любой момент времени.

Литература

1. *Erich, A.* Techniques of Financial Analysis with Financial Genome Passcode Card.- *McGraw-Hill/Irwin*, 2002, 570 p.
2. *Esbensen, K. , Guyot, D., Westad, F.* Multivariate Data Analysis - in Practice. Multivariate Data Analysis, 2002, 598 p
3. *Fein, M. L.*, Robo-advisors: A closer look, 2015
4. *Markowitz, H.*, PORTFOLIO SELECTION. *Journal of Finance*, 7, issue 1, 1952, p. 77-91
5. *Marvin, K.* Creating Diversified Portfolios Using Cluster Analysis. Independent Work Report Fall, 2015
6. *Melanie L. Fein.* Robo-advisors: a closer look 2015
7. *Ren, Z.* Portfolio construction using clustering methods (Doctoral dissertation, Worcester Polytechnic Institute). Retrieved from <https://web.wpi.edu/Pubs/ETD/Available/etd-042605-092010>, 2005
8. *Тюкхова, Е., Сизыкх, Д., & Smirnov, A.* (2018). Quality Estimation Model of Investment Portfolio Rebalancing Process. Proceedings of 2018 Eleventh International Conference "Management of large-scale system development", pp. 1-5. IEEE. Retrieved from <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8551950>