

УДК 330.322.01

EDN [XHACUZ](#)



Использование метода кластерного анализа для анализа данных фондового рынка

А.А. Терешенко*

ФГАОУ ВО «Северный (Арктический) федеральный университет имени М. В. Ломоносова», Архангельск, Россия

*E-mail: andrey.tereshenko2017@mail.ru

Аннотация. В данной статье рассматривается возможность использования метода кластерного анализа для анализа больших данных фондового рынка. На сегодняшний день очень востребовано и актуально применение методов машинного обучения в различных областях. В статье дана краткая характеристика метода кластерного анализа. Отмечено, почему может появиться необходимость в его использовании. В частности, это связано с большим количеством финансовых инструментов, что может усложнить создание инвестиционного портфеля. Также отмечены положительные стороны использования кластерного анализа, такие как: грамотный подход к кластерной торговле, создание диверсифицированного портфеля, устранения дублирования в портфелях. Недостатком является: уровень перекрытия между кластерами. В конце представлены выводы, доказывающие важность и перспективы применения рассматриваемого метода.

Ключевые слова: машинное обучение, кластерный анализ, диверсификация.

Using the cluster analysis method to analyze stock market data

A.A. Tereshenko*

Northern (Arctic) Federal University, Arkhangelsk, Russia

*E-mail: andrey.tereshenko2017@mail.ru

Abstract. This article discusses the possibility of using the cluster analysis method for analyzing big stock market data. Today, the application of machine learning methods in various fields is very in demand and relevant. The article gives a brief description of the cluster analysis method. It is noted why it may be necessary to use it. In particular, this is due to a large number of financial instruments, which can complicate the creation of an investment portfolio. The positive aspects of using cluster analysis are also noted, such as: a competent approach to cluster trading, the creation of a diversified portfolio, and the elimination of duplication in portfolios. The disadvantage is: the level of overlap between clusters. At the end of presentations proving significant and promising applications of the proposed method.

Keywords: machine learning, cluster analysis, diversification.

1. Введение

В эпоху информации и цифровой трансформации область искусственного интеллекта, наряду со своими подкатегориями, машинным обучением и глубоким обучением является очень востребованной и актуальной. По данным различных исследований, предполагается, что искусственный интеллект станет самым большим прорывом в большинстве отраслях в ближайшие пять лет. Искусственный интеллект влияет на многие основы, обеспечивающие цифровую трансформацию.

Машинное обучение — это дисциплина искусственного интеллекта, которая предоставляет машинам возможность автоматически учиться на данных и прошлом опыте, выявляя закономерности и делая прогнозы с минимальным вмешательством человека [1].

Получая новые данные, приложения машинного обучения могут самостоятельно учиться, расти, развиваться и адаптироваться. Машинное обучение позволяет извлекать полезную информацию из больших объемов данных, используя алгоритмы для выявления закономерностей и обучения в итеративном процессе [2]. Алгоритмы машинного обучения используют методы вычислений, чтобы учиться непосредственно на данных вместо того, чтобы полагаться на какое-либо заранее определенное уравнение, которое может служить моделью.

2. Материалы и методы

В данной статье рассматривается кластерный анализ. Это статистический метод обработки данных. Он работает путем организации элементов в группы или кластеры на основе того, насколько тесно они связаны. Кластерный анализ, касается матриц данных, в которых переменные не были заранее разделены на подмножества критериев и предикторов. Цель кластерного анализа состоит в том, чтобы найти похожие группы субъектов, где «сходство» между каждой парой субъектов означает некоторую глобальную меру по всему набору характеристик.

Задачей данной статьи является изучение возможностей применения метода кластерного анализа для анализа данных фондового рынка. Изначально будет обоснована необходимость использования данного метода машинного обучения на фондовом рынке. Далее будут рассмотрены плюсы и недостатки его использования. В

конец сделаны выводы, а также описаны дальнейшие перспективы данного метода относительно фондового рынка.

Кластерный анализ — это алгоритм обучения без учителя, означающий, что изначально неизвестно, сколько кластеров существует в данных до запуска модели. В отличие от многих других статистических методов, кластерный анализ обычно используется, когда нет никаких предположений о вероятных взаимосвязях между данными. Он предоставляет информацию о том, где существуют ассоциации и закономерности в данных, но не о том, что это может быть или что они означают.

3. Результаты и обсуждение

В современном мире наблюдается рост количества инвестиционных проектов. С каждым днём увеличивается набор инвестиционных возможностей, что создаёт сложность при сборе инвестиционного портфеля. Это является проблемой.

Поскольку объём финансовых рынков расширился с появлением инновационных финансовых инструментов и ценных бумаг, а также различных производных продуктов, для инвесторов крайне важно понять тенденцию структурного совместного движения инвестиционной вселенной и оценить набор инвестиционных возможностей. Для решения данной задачи подойдет метод кластерного анализа.

С помощью кластерного анализа инвестор сможет разработать подход к кластерной торговле, благодаря которому можно будет создать диверсифицированный портфель. Как вариант, акции с высокой корреляцией доходности попадают в одну корзину, акции с меньшей корреляцией — в другую и так далее, пока каждая акция не будет помещена в определенную категорию [3].

При грамотном подходе разные кластеры будут демонстрировать минимальную корреляцию друг с другом. Таким образом, инвестор получит все преимущества диверсификации: снижение убытков, сохранение капитала и возможность совершать более рискованные сделки, не увеличивая общий риск [4]. Диверсификация остается одним из главных факторов инвестирования, а кластерный анализ — лишь один из способов ее достижения.

Кластерный анализ может позволить инвестору устранить дублирование в своих портфелях, определяя ценные бумаги с соответствующей доходностью. Рассмотрим пример. У инвестора имеется портфель, состоящий только из технологических акций,

может показаться безопасным и диверсифицированным на первый взгляд, но, когда происходит событие, подобное пузырю доткомов, оказывается, что портфель становится очень уязвимым – как итог инвестор получит большие потери. Покупка и объединение активов, подходящих для разных сегментов рынка, имеет решающее значение для увеличения диверсификации и защиты от таких системных рисков

Этот метод также может выявить определенные категории акций, такие как циклические акции и акции роста. Эти конкретные стратегии подпадают под умную бета-версию или факторное инвестирование. Они пытаются получить более высокую доходность с поправкой на риск от конкретных надбавок за риск, таких как минимальная волатильность, рост и импульс [5]. В некотором смысле умное бета- или факторное инвестирование воплощает концепции группировки и категоризации, проповедуемые кластерным анализом. Логика кластеризации по одному общему поведению отражает базовую методологию факторного инвестирования, которая определяет акции, подверженные схожим системным рискам и имеющие схожие характеристики. Не всегда активы в кластере относятся к одной отрасли. Часто кластеры содержат акции из нескольких отраслей, таких как технологии и финансы.

Очевидным недостатком кластерного анализа является уровень перекрытия между кластерами. Кластеры близки по расстоянию, что означает высокую корреляцию доходности, часто имеют некоторые схожие факторы риска. Таким образом, день простоя в одном кластере может привести к такой же низкой производительности в другом кластере. По этой причине инвестору следует находить и группировать акции с большим расстоянием между ними. Таким образом, на кластеры влияют различные рыночные факторы. Тем не менее, масштабные откаты рынка, такие как рецессия 2008 года, затормозят весь портфель независимо от его структуры. Даже самые диверсифицированные кластеры не смогут противостоять встречным ветрам рецессии. Здесь лучшее, что может сделать кластеризация, — это свести к минимуму экстремальные потери.

4. Заключение

В заключении подведём итоги. На сегодняшний день существуют высокие темпы развития искусственного интеллекта, в частности машинного обучения. Используя метод кластерного анализа, появляется возможность: диверсифицировать свой

портфель, исключить дублирование в нём каких-либо акций, а также с его помощью можно выделить циклические акции и акции роста в своём портфеле. Очевидным недостатком является уровень перекрытия между кластерами, то есть кластеры близки по расстоянию, что означает высокую корреляцию доходности. Также важно отметить зависимость от случайных событий, которые невозможно предсказать, например кризис 2008 года. Учитывая, как быстро и точно искусственный интеллект приходит в нашу жизнь, нельзя исключать, что крупные инвестиционные компании, частные инвесторы прибегнут к использованию искусственного интеллекта в своей торговле. На сегодняшний день, никак нельзя рассмотреть случай, где инвестора или трейдера сможет заменить искусственный интеллект. Однако, если рассматривать вариант его использования как вспомогательный инструмент, то такой вариант возможен.

Список литературы

1. Машинное обучение: просто о сложном: [сайт]. – URL: <https://sbercloud.ru/ru/warp/blog/machine-learning-about> (дата обращения: 01.04.2023). – Текст: электронный.
2. Машинное обучение – основа искусственного интеллекта: [сайт]. – URL: https://habr.com/ru/hub/machine_learning/ (дата обращения: 01.04.2023). – Текст: электронный.
3. Грэм Б. Разумный инвестор: Полное руководство по стоимостному инвестированию / Бенджамин Грэм; Пер. с англ. – 5-е изд. – 2021. – 568 с.
4. Миллер Дж. Правила инвестирования Уоррена Баффета / Джереми Миллер; Пер. с англ. В. Ионова. – 2-е изд. – 2021. – 372 с.
5. Нисон С. Японские свечи: Графический анализ финансовых рынков / Стив Нисон, Пер. с англ. – 3-е изд. 2020. – 290 с.