

Сухих Юрий Владимирович, Магистрант,
ФГБОУ ВО «Вятский Государственный Университет»,
Киров

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ПОДБОРА ТРЕЙДЕРОВ ДЛЯ КОЛЛЕКТИВНОЙ ТОРГОВЛИ НА ВЫСОКОВОЛАТИЛЬНЫХ РЫНКАХ

Аннотация: В данной статье будет представлен способ подбора трейдеров для инвестора на рынках с высокой волатильностью при помощи гибридной рекомендательной системы. Будет предложен собственный вариант реализации гибридной рекомендательной системы, позволяющий при помощи набора продукционных правил объединить результаты разных систем.

Ключевые слова: Инвестиции, коллективное следование, оценка рисков, высоковолатильные рынки, трейдинг, рекомендательные системы.

Для начала рассмотрим предметную область, на которую направлена данная статья, для большего понимания идеи предложенных в ней.

Рекомендательная система – это программная система, предназначенная для прогнозирования предпочтений пользователя на основе анализа его поведения [1] По большей части задачей таких систем является определение полезного в текущем контексте объекта для неопределенного пользователя. Далее рассмотрим три основных вида рекомендательных системы: коллаборативная фильтрация, контент ориентированные, гибридные. Теперь подробнее изучим каждую из них.

Коллаборативная фильтрация – это метод отбора контента, основанный на мнениях других пользователей системы, который похож на рекомендацию чего-либо знакомым. Этот подход достаточно широко изучен и применяется, например, в интернет-магазинах для персонализации рекомендаций. В работе коллаборативной фильтрации используются матрицы интересов, при помощи которых можно определить какие предметы интересуют каждого пользователя. Эти матрицы позволяют системе находить пользователей с похожими интересами и делать рекомендации на основе их предпочтений.

Одной из важнейших частей работы систем коллаборативной фильтрации является метрика, при помощи которой они определяют схожесть предметов или пользователей между собой. Далее рассмотрим одну из таких метрик [2].

Коэффициент корреляции Пирсона. Один из популярных методов в коллаборативной фильтрации, используемый для линейной корреляции между двумя векторами [2] Данный метод определяет, насколько одна последовательность чисел изменяется относительно другой. Чем сильнее вектора согласованы, тем ближе значение коэффициента к 1. Если вектора слабо связаны между собой, коэффициент будет близок к 0. Вычисление данного коэффициента производится по формуле

$$\text{sim}(i, j) = \frac{\langle R_{k,i} - A_i, R_{k,j} - A_j \rangle}{\|R_{k,i} - A_i\| \|R_{k,j} - A_j\|} = \frac{\sum_{k=1}^n (R_{k,i} - A_i) (R_{k,j} - A_j)}{\sqrt{\sum_{k=1}^n (R_{k,i} - A_i)^2 * \sum_{k=1}^n (R_{k,j} - A_j)^2}} \quad (1)$$

где $R_{k,i}$ – рейтинг предмета i от пользователя k ;

$R_{k,j}$ – рейтинг предмета j от пользователя k ;

A_i – средний рейтинг предмета i ;

A_j – средний рейтинг предмета j ;

n – общее число оценок пользователей для предметов.

Системы, ориентированные на пользователя, сталкиваются с проблемой поиска "идеальных соседей". Поиск ближайших пользователей может быть ресурсоемким, поэтому как правило выделяется меньшая группа пользователей или проводится кластеризация, чтобы сравнивать пользователей внутри кластеров. Однако такие методы также могут привести к потере точности модели.



Контент-ориентированные системы являются другим видом рекомендательных системы. Контент-ориентированные рекомендательные системы представляют собой альтернативный подход к рекомендациям, отличный от систем коллаборативной фильтрации. Вместо того чтобы полагаться на оценки других пользователей, они анализируют содержимое самих объектов. Этот метод исследует характеристики объектов и сравнивает их между собой по содержимому, полученному на основе выделения ключевых характеристик [3] Например, если k_i – является i -м атрибутом предмета d_j , то w_{ij} является весом k_i для d_j . Тогда содержимое объекта можно описать следующим образом.

$$Content(d_j) = \{w_{1j}, w_{2j}, \dots\} \quad (2)$$

По итогу можно создать векторную функцию, которая на основе содержания предмета и профиля интересов пользователя оценит, на сколько именно этот предмет заинтересует конкретного пользователя.

Степень схожести документов вычисляется по следующей формуле [4].

$$sim(d_i, d_j) = \frac{\sum_k w_{ki} * w_{kj}}{\sqrt{\sum_k w_{ki}^2} * \sqrt{\sum_k w_{kj}^2}} \quad (3)$$

Способность определить схожесть между предметами без учета оценок даёт возможность преодолеть проблему "холодного старта" в рекомендательных системах. На старте системы сложно определить, кому рекомендовать какие-либо предметы, так как информация о их сходстве с другими отсутствует.

Гибридные системы

Для исправления недостатков рассмотренных ранее систем был разработан класс гибридных систем. Эти системы объединяют несколько методов в один, чтобы уменьшить или и вовсе устранить их недостатки по отдельности. Для такого подхода существует несколько способов комбинирования рекомендательных систем, [5] рассмотрим некоторые из них:

- взвешенная комбинация. Здесь используются несколько рекомендательных систем, каждой из которых присваивается определенный вес. Результатом является линейная комбинация рейтингов от каждой системы.;

- переключение. Система переключается между различными методами рекомендаций в зависимости от контекста. Например, в начале работы пользователя в системе ему могут предлагаться популярные товары, а затем будут уже более персонализированные рекомендации;

- конвейер. Методы используются последовательно, например, для фильтрации нерелевантных сущностей перед передачей результатов другой системе для дальнейшей обработки.

Далее рассмотрим работу модуля рекомендаций. Диаграмма компонента модуля рекомендаций представлена на рисунке 1.

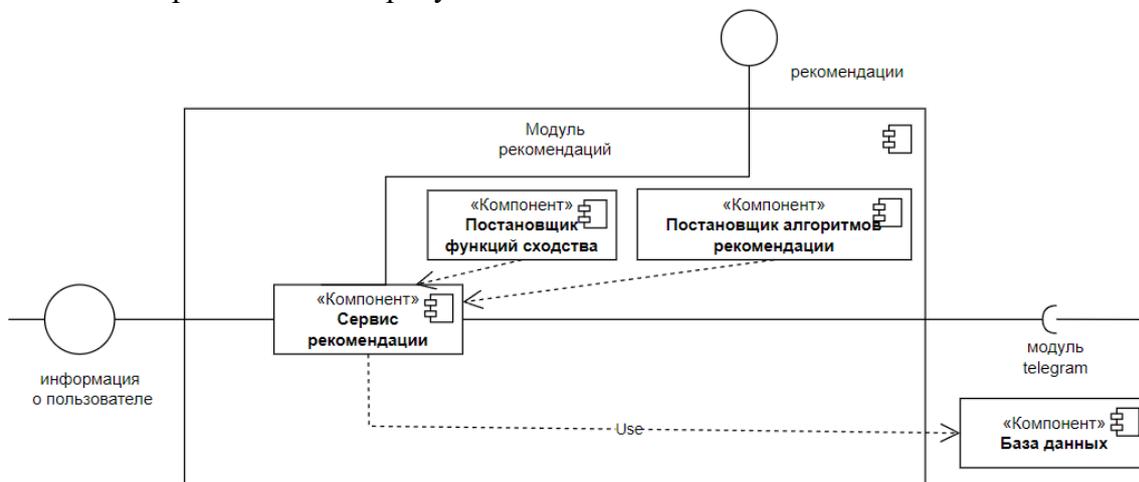


Рисунок 1 – Диаграмма компонента модуля рекомендаций



В зону ответственности данного модуля входит анализ и расчет рекомендаций трейдеров для подписчиков. В его задачи входит генерация персонализированных рекомендаций на основе предпочтений и прошлой активности пользователя, запрос рекомендаций для подписчиков, а также обработка процессов добавления новых оценок. В случае появления новых оценок трейдеров, модуль должен инициировать фоновые задачи по обновлению рекомендаций, чтобы обеспечить актуальность и точность рекомендаций для подписчиков. Этот модуль играет важную роль в повышении эффективности стратегий копирования сделок для подписчиков, предлагая им наилучшие варианты трейдеров в соответствии с их потребностями и целями инвестирования.

Далее рассмотрим алгоритмы данного модуля. Данный модуль отвечает за рекомендацию подходящих пользователю трейдеров. При запросе пользователем списка трейдеров будут получены как вектор оценки пользователя, так и вектора всех трейдеров, из которых будут выбраны те, что совпадают с предпочтениями пользователя и будут показаны ему отсортированными по убыванию оценки. Однако перед этим по завершению работы рекомендательных систем будет необходимо произвести расчеты, которые позволят учесть в итоговой рекомендации результат работы обеих систем.

На рисунке 2 представлена схема алгоритма модуля рекомендаций.

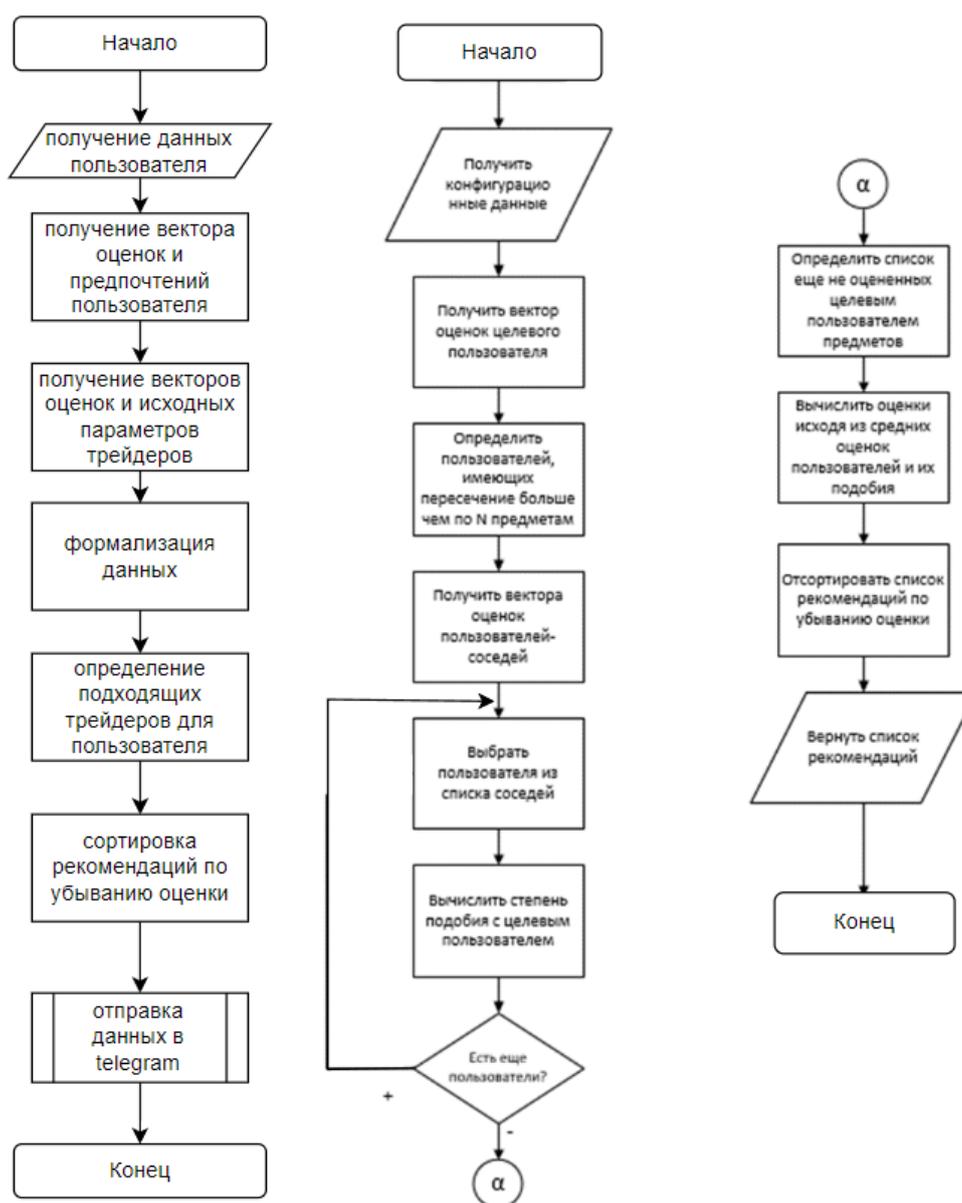


Рисунок 2 – Схема алгоритма работы модуля рекомендаций



В качестве метода совмещения результатов работ контент ориентированной системы и системы коллаборативной фильтрации был выбран способ взвешенной комбинации, в основе которого будет лежать набор продукционных правил, использование которых позволит более грамотно оценить влияние каждой из систем на итоговый результат. Далее рассмотрим одно из таких правил, оно уменьшает вес результата коллаборативной фильтрации практически до нуля, при недавней регистрации и отсутствии длительных подписок. Таким образом изначально система в начале будет полагаться на интересы, указанные пользователем, позволяя медленно подстроиться под его реальные интересы не уменьшая качество рекомендаций. Сам итоговый результат работы гибридной рекомендательной системы будет рассчитан следующим образом:

$$R = \alpha \cdot R_c + \beta \cdot R_f \quad (4)$$

где R_c – рейтинг полученный от контент ориентированной системы;

R_f – рейтинг, полученный от системы коллаборативной фильтрации;

α – коэффициент, отражающий вес контент ориентированной системы;

β – коэффициент, отражающие вес системы коллаборативной фильтрации.

В заключение следует сказать, что практическая ценность данной работы заключается в том, что такой вариант реализации системы позволит инвестору указать наиболее подходящие для него параметры торговли, на основе которых порекомендует наиболее подходящих ему трейдеров, что по итогу позволит инвесторам значительно сократить свои временные затраты или же и вовсе избавит от них в некоторых моментах.

Список литературы:

1. Меньшикова Надежда Васильевна, Портнов Иван Владимирович, Николаев Иван Евгеньевич Обзор рекомендательных систем и возможностей учета контекста при формировании индивидуальных рекомендаций // Academy. 2016. №6 (9)
2. Yunkyoungh, L. RECOMMENDATION SYSTEM USING COLLABORATIVE FILTERING / L. Yunkyoungh – Текст: непосредственный // San Jose State University – 2015. – С. 1-49.
3. Ma, K. Content-based Recommender System for MovieWebsite / K. Ma. – Stockholm, Sverige, 2016. – 50с.
4. Lops, P. Content-based Recommender Systems: State ofthe Art and Trends / P. Lops, M. de Gemmis, G. Semerano. – Текст: непосредственный // Recommender Systems Handbook. – 2011. – С. 74-105.
5. Liu A., Calvik J., Using Demographic Information to Reduce the New User Problem in Recommender Systems. – Sweden: 2017. – 35 с

