

АНАЛИТИЧЕСКИЙ ОБЗОР ПРИМЕНЕНИЯ ТЕХНОЛОГИЙ МАШИННОГО ОБУЧЕНИЯ В ФИНАНСОВЫХ АССИСТЕНТАХ

Пилявская И.М.

*Пилявская Ирина Михайловна - руководитель аналитики,
Тинькофф Банк, г. Москва*

Аннотация: целью данного аналитического обзора технологий искусственного интеллекта является выбор наиболее подходящего метода машинного обучения для построения рекомендаций инструментов для управления личными финансами пользователей финансовых ассистентов. С помощью данного исследования были рассмотрены особенности методов машинного обучения и аргументирован выбор гибридного вида рекомендательной системы.

Ключевые слова: финансовый ассистент, финансовое планирование, управление инвестиционным портфелем, технологии машинного обучения.

ANALYTICAL REVIEW OF THE APPLICATION OF MACHINE LEARNING TECHNOLOGIES IN FINANCIAL ASSISTANTS

Pilyavskaya I.M.

*Pilyavskaya Irina Mikhailovna - Head of Analytics,
TINKOFF BANK, MOSCOW*

Abstract: the purpose of this analytical review of artificial intelligence technologies is to select the most appropriate machine learning method for building recommendations for personal finance management tools for users of financial assistants. With the help of this study, the features of machine learning methods were considered and the choice of a hybrid type of recommender system was argued.

Keywords: financial assistant, financial planning, investment portfolio management, machine-learning technologies.

УДК 004.42

Искусственный интеллект (ИИ) уже не является чем-то удивительным в приложениях для управления личными финансами. Современные компании, занимающиеся разработкой финансового программного обеспечения, и банковские компании расширяют свои услуги, разрабатывая помощников по личным финансам на основе искусственного интеллекта (рис. 1).

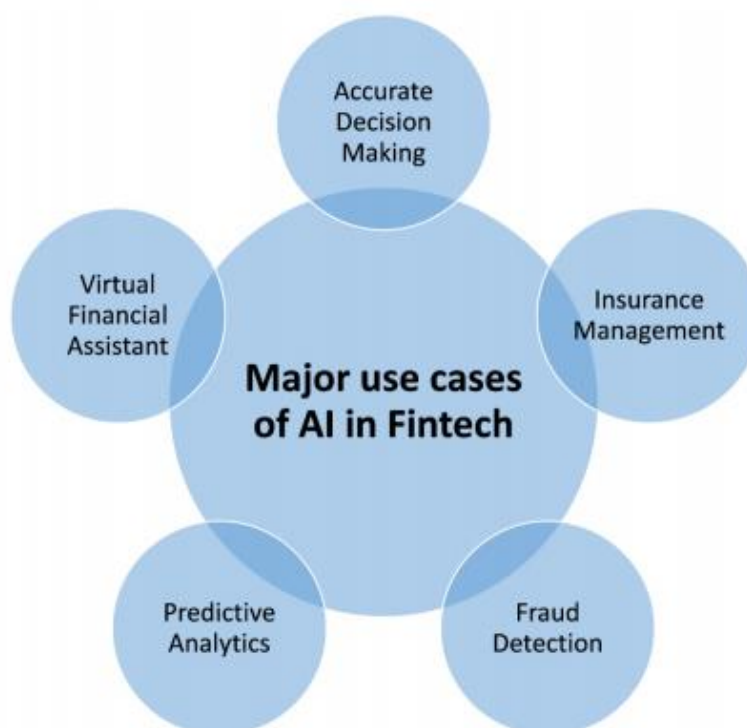


Рис. 1. Основные задачи ИИ в сфере финансовых технологий

Разработка приложения для личных финансов на базе искусственного интеллекта — это способ для компаний финансового сектора помочь своим клиентам принимать правильные финансовые решения. Согласно исследованию NerdWallet, 49% американцев тратят на эмоциях больше, чем они могут себе позволить. Благодаря рекомендациям ИИ пользователи могут избавиться от вредных привычек, тратить и лучше управлять своими деньгами.

Еще одна причина, по которой компании создают помощников на основе ИИ для финансирования — это наладить личные отношения со своими клиентами. Помощник ИИ общается быстрее, чем люди, и доступен в любое время, поэтому клиентам проще общаться с виртуальными помощниками, чем с настоящими. Другими словами, ИИ обеспечивает более приятный пользовательский опыт, чем традиционные методы коммуникации.

Поддержка принятия решений в большинстве виртуальных ассистентов осуществляется на основании стандартных рекомендательных алгоритмов.

Рекомендательная система представляет собой программное средство и методы, назначением которых является прогнозирование поведения пользователя в отношении объекта информационного поиска и формирование рекомендаций для объектов, с которыми он еще не встречался. Потребность в персонализации информационных предложений также будет реализовываться в развитие интеллектуальных советников, которые будут адаптироваться под конкретного пользователя. Источником данных для рекомендательной системы выступают профили пользователей. В них фиксируются данные о самом пользователе, предоставленные как лично, так и через перечень действий, которые он совершил за время работы с системой.

Формальная постановка задачи для рекомендаций выглядит следующим образом. Рассмотрим U — множество пользователей и D — множество объектов.

Необходимо найти функцию $r, r: U \times D \rightarrow R$, которая формирует рекомендацию R таким образом, что для любого пользователя значение r между ним и объектом i максимально, т.е. является аргументом максимизации (1):

$$\forall u \in U, d_u = \arg(\max_{d \in D} r(u, d)). \quad (1)$$

Рекомендательные системы можно классифицировать как показано на рис. 2.

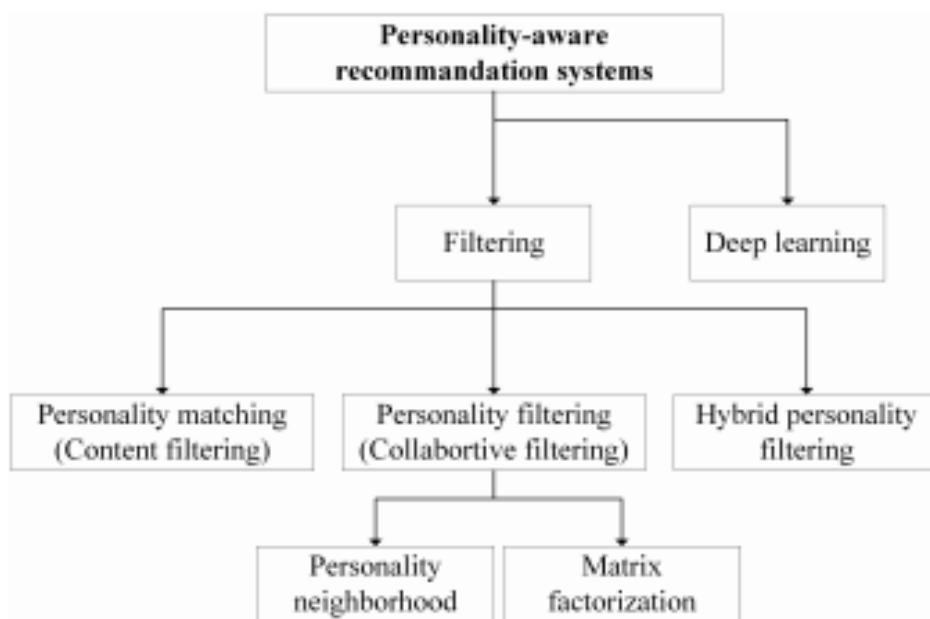


Рис. 2. Классификация рекомендательных систем

Выделяют четыре вида рекомендательных систем:

- рекомендации, формируемые экспертным методом – связи между объектами устанавливаются вручную или на основе заранее определенных правил, но данный способ является актуальным только при небольшом перечне рекомендуемых объектов;

- коллаборативная фильтрация – рекомендации, которые основываются на оценках пользователей по отношению к просмотренным объектам, то есть на том, что понравилось пользователю. Рекомендации могут строиться либо на основе поиска схожих пользователей по отношению к рассматриваемому пользователю (user-based), либо на основе поиска схожих объектов по отношению к объектам с выставленными ранее оценками рассматриваемым пользователем (item-based);

– контентная фильтрация – рекомендации для рассматриваемого пользователя, которые формируются на основе понравившихся ему объектов, то есть основываясь только на знаниях о схожести контента с учетом присвоенных каждому объекту набору параметров без персонализированных рекомендаций;

– гибридная фильтрация – используется комбинация подходов, основанных на контентной и коллаборативной фильтрации, где оценивается как схожесть пользователей, так и схожесть объектов. Это приводит к повышению качества формируемых рекомендаций.

Персонализированная система рекомендаций финансового ассистента должна учитывать не только пол, возраст, основное занятие и состав семьи, но также финансовую поведенческую модель человека. Также было рассмотрено исследование о рекомендательных системах, основанных на профилях пользователей по интересам.

Рассмотрим наиболее успешные и потому наиболее популярные из применяемых методов.

Метод коллаборативной фильтрации ориентирован на сравнение интересов пользователей между собой и поиск пользователей со схожими интересами. Данный метод использует информацию о поведении пользователей в прошлом, например: информация о запросах, оценках и т.п. В этом случае не имеет значения, с какими типами объектов ведётся работа, но при этом могут учитываться неявные характеристики, которые сложно было бы учесть при создании профиля. Основная проблема этого типа рекомендательных систем – отсутствие данных о недавно появившихся в системе пользователях или объектах, так как в этом случае отсутствует предыстория.

Для определения схожести пользователей выполняется расчет близости векторов оценок пользователей, который может быть выполнен с помощью вычисления косинуса между соответствующими векторами (2):

$$sim(x, y) = \frac{\sum_{d=1}^m v_{x,d} \cdot v_{y,d}}{\sqrt{\sum_{d=1}^m v_{x,d}^2} \cdot \sqrt{\sum_{d=1}^m v_{y,d}^2}}, \quad (2)$$

где: $sim(x, y)$ – мера близости между пользователями x и y ;

$v_{x,d}, v_{y,d}$ – оценка пользователей x и y для объекта d из матрицы V .

Контентная фильтрация направлена на четкую классификацию как пользователей, так и объектов, фигурирующих в информационном предложении. В указанном случае устанавливается прямое соответствие между пользователями и объектами на основе их характеристик. В целом стратегия хорошо работает в областях с конечным и относительно небольшим количеством критериев оценки, вытекающих из природы вещей при больших потоках информации, и допускает большое количество критериев при небольшом информационном потоке.

При фильтрации на основе контента (содержания) создаются профили пользователей и объектов: профили пользователей могут включать демографическую информацию или ответы на определённый набор вопросов; профили объектов – различную атрибутивную информацию в зависимости от типа объекта.

В случае контентной фильтрации рекомендации даются для объектов, похожих на объекты, уже выбранные пользователем, а основная идея сводится к тому, что если пользователю понравился один объект, то ему должен понравиться и аналогичный. Контентная фильтрация изначально была разработана для формирования текстовых рекомендаций, которые описывают текстовый объект определенным набором термов. Для формирования рекомендаций из такого набора данных применяют специальную метрику TF-IDF, которая позволяет определить важность терма в рамках конкретного предложения и, тем самым, повысить значимость данного текстового документа по сравнению с другими документами.

Формальная постановка задачи для контентных рекомендательных систем

выглядит следующим образом. Рассмотрим множество текстовых документов D .

Пусть $P(d_j)$ будет являться профилем документа $d \in D$, представляющее собой множество признаков данного объекта. Каждый признак представляет собой коэффициент важности терма в документе d , рассчитанный в соответствии с метрикой TF-IDF и имеет следующий вид (3):

$$\vec{P}(d_j) = (w_{1j}, w_{2j}, \dots, w_{kj}), \quad (3)$$

где: $P(d_j)$ – профиль документа, представленный вектором взвешенных термов (слов) из значений термов;

d_j – текстовый документ;

w_{kj} – вес терма, рассчитанный с помощью метрики TF-IDF.

Рекомендация строится на основании оценки близости контента профиля пользователя и контента объекта (4):

$$r(u, d) = Score(P(u), P(d)), \quad (4)$$

где: $r(u, d)$ – мера близости;

$P(u)$ – профиль пользователя, представленный вектором взвешенных термов (слов) на основе просмотренных пользователем объектов;

$P(d)$ – профиль объекта, представленный вектором взвешенных термов (слов).

Для построения рекомендации финансовых инструментов для решения задач управления личными финансами пользователей необходимо выдавать список заданного числа наиболее подходящих к рекомендации объектов (финансовых инструментов). В таком случае необходимо использовать сразу все доступные данные:

- данные о финансовом портрете пользователя;
- данные о финансовом инструменте;
- данные пользовательских оценок.

Таким образом, для решения поставленной задачи наиболее подходит гибридный вид рекомендательной системы, которая позволит учесть все доступные данные и избавиться от ряда проблем, например, от проблемы «холодного старта».

Список литературы / References

1. Intelligent Virtual Assistant Market Size, Share & Trends Analysis Report By Product (Chatbot, Smart Speakers), By Technology, By Application (BFSI, Healthcare, Education), By Region, And Segment Forecasts, 2020 – 2027 Grand View Research. 4/1/2020. 130 Pages - ID: GV16217098. [Электронный ресурс]. Режим доступа: <https://www.marketresearch.com/Grand-View-Research-v4060/Intelligent-Virtual-Assistant-Size-Share-13187162/>
2. *Markets R.A.* Intelligent Virtual Assistant (IVA) Market - Growth, Trends, and Forecast (2020 - 2025) [Электронный ресурс]. Режим доступа: <https://www.researchandmarkets.com/reports/4845914/intelligent-virtualassistant-iva-market/> (дата обращения: 05.02.2021).
3. Statista – Personal Finance. [Электронный ресурс]. Режим доступа: <https://www.statista.com/outlook/dmo/fintech/personal-finance/worldwide/> (дата обращения: 05.01.2021).
4. Personal Finance Topics: Best Personal Finance Apps. [Электронный ресурс]. Режим доступа: <https://www.thebalance.com/best-personal-finance-apps-4170650/> (дата обращения: 10.01.2021).
5. *Olivia A.I.* [Электронный ресурс]. Режим доступа: <https://www.olivia.ai/> (дата обращения: 22.01.2021).
6. You Need a Budget Features. [Электронный ресурс]. Режим доступа: <https://www.youneedabudget.com/features/> (дата обращения: 22.01.2021).
7. Personal Capital Wealth Management. [Электронный ресурс]. Режим доступа: <https://www.personalcapital.com/wealth-management/> (дата обращения: 22.01.2021).
8. Prism by BillGo Product Description. [Электронный ресурс]. Режим доступа: <https://www.prismmoney.com/> (дата обращения: 22.01.2021).
9. Hello Digit, Inc. [Электронный ресурс]. Режим доступа: <https://digit.co/> (дата обращения: 22.01.2021).
10. Spendee Pricing [Электронный ресурс]. Режим доступа: <https://www.spendee.com/pricing/> (дата обращения: 22.01.2021).
11. EveryDollar Product Description. [Электронный ресурс]. Режим доступа: <https://www.everydollar.com/> (дата обращения: 22.01.2021).
12. Royal Bank of Canada Website. [Электронный ресурс]. Режим доступа: <https://www.rbcroyalbank.com/mobile/feature/nomi/index.html/> (дата обращения: 22.01.2021).
13. *Hornýák A.*, 2015. Attitudes and Competencies Among High School Students Regarded as Potential Banking Clients. PhD Thesis. Sopron: University of West-Hungary, Faculty of Economics.
14. *Rosenberg M.J.*, 1960. An analysis of affective-cognitive consistency. In: M.J. Rosenberg, C.I. Hovland, W.J. McGuire, R.P. Abelson, J.W. Brehm (ed.): Attitude organization and change. New Haven: Yale University Press.
15. *Goldberg H., Lewis R.*, 1978. Money Madness. London: Springwood Books.