



**Главный
радиочастотный
центр**

**НАУЧНО-ТЕХНИЧЕСКИЙ ЦЕНТР
ФГУП «ГРЧЦ»**

**АНАЛИТИЧЕСКИЙ ОБЗОР
«РЕКОМЕНДАТЕЛЬНЫЕ СЕРВИСЫ:
МЕТОДЫ И РЕГУЛИРОВАНИЕ»**

Москва
2023



ОГЛАВЛЕНИЕ

Аннотация	3
Определение рекомендательных технологий и сервисов	7
Рынок рекомендательных технологий	8
Типы рекомендательных сервисов и используемые методы	8
Рекомендательные алгоритмы в социальных сетях	12
Законодательство в отношении рекомендательных сервисов	14
Китай	14
Европейский Союз	16
США	17

АННОТАЦИЯ

В условиях глубокого проникновения информационных технологий в жизнь людей большую актуальность приобретают вопросы регулирования рекомендательных сервисов. Социальные сети и интернет-приложения активно используют рекомендательные системы для индивидуализированного подбора и ранжирования контента для пользователей.

Мировой рынок рекомендательных технологий в 2020 г. был оценен в 132,5 млрд руб., к 2028 г. данный показатель должен достигнуть значения в 1 трлн 297 млрд руб. Объем российского рынка в 2020 г. оценивался в 6 млрд руб., к 2024 г. прогнозируется его рост до 21,7 млрд руб.

Определение и методы

Рекомендательные системы и сервисы — это программные комплексы, которые с помощью алгоритмических вычислений и машинного обучения на основании данных о пользователе или характеристиках элементов в системе осуществляют индивидуализированный подбор, а также ранжирование контента или товаров для конечного пользователя.

Основные методы, применяемые в рекомендательных системах, — *фильтрация на основе контента* и *коллаборативная фильтрация*.

При *фильтрации на основе контента* алгоритмы рекомендуют товары или контент, похожие на те, которые пользователю нравились в прошлом или которые он изучает в настоящее время.

При *коллаборативной фильтрации* используется информация о поведении пользователей с похожими интересами. Система находит пользователей или элементы с историей оценок, аналогичной текущему пользователю или элементу, и генерирует рекомендации на основании этой схожести.

Большинство рекомендательных сервисов используют гибридный подход, объединяя эти две технологии. Именно такой подход дает наиболее точный результат.

Использование искусственного интеллекта (ИИ) в таргетированной рекламе

Алгоритмы позволяют предсказать, будет ли реклама эффективна в социальных сетях еще до того, как она запущена, заранее оптимизировать объявление и получить максимум переходов по ссылке. Алгоритмы способны управлять рекламными кампаниями целиком, корректируя настройки таргетинга, креативы и бюджет в режиме реального времени.

Некоторые соцсети самостоятельно осуществляют премодерацию политической рекламы. Так, корпорация Meta проводит процедуру получения разрешения на размещение политической рекламы в Facebook¹: рекламодатель должен проживать или находиться в стране на время рекламной кампании, а в объявлениях должна присутствовать информация об источнике финансирования с указанием организации, оплатившей рекламу, и имени ответственного лица. В противном случае политическая реклама будет приостановлена, а разрешение на показ — отозвано. Вместе с тем корпорация открыто заявляла о возможности публикации призывов к насилию по отношению к русским на территории Украины, т. е. модерация выборочна.

¹ Признана экстремистской организацией и запрещена в России.

Персонализированная новостная лента

Алгоритм анализирует взаимодействие пользователя с платформой и прогнозирует, с каким контентом он будет взаимодействовать больше всего.

Так, ВКонтакте использует рекомендательные сервисы для формирования ленты пользователя, но функцию можно отключить и просматривать новости в хронологическом порядке. Выбор контента в ленте Дзен основан на анализе истории посещенных страниц, пользовательских предпочтениях, местоположении, времени суток и других факторах.

При распространении контента среди пользователей нейросеть часто отдает предпочтение спорному, разжигающему вражду и недостоверному контенту, поскольку именно такой контент привлекает больше внимания пользователей и ведет к повышению вовлеченности во взаимодействие с платформой. Исследователи неоднократно отмечали эту закономерность² в отношении работы Facebook AI³. Алгоритм, нацеленный на распространение контента с максимальной вовлеченностью, в итоге продвигал среди пользователей огромное количество фейковых новостей и риторику ненависти. Эту проблему, по мнению экспертов, можно решить, только вручную фильтруя контент и изменяя алгоритм.

Предложение возможных «друзей» и «подписок»

Рекомендации формируются на основе структурной схожести профилей (например, большое количество общих друзей), что приводит к повышению коэффициента кластеризации онлайн-платформы, т. е. группы пользователей становятся все более изолированными.

Другим следствием использования рекомендательных алгоритмов может быть создание «эхокамер» — окружение пользователя информацией схожего характера. Это происходит, когда рекомендательная система предлагает пользователю контент других пользователей на основании схожести их взглядов и интересов. Человек оказывается огражден от информации другого типа или от критической информации в отношении собственных взглядов.

В результате социальные сети могут спровоцировать политическую поляризацию и продвигать дезинформацию. Так, 64% всех новых подписчиков экстремистских групп пришли из рекомендаций.

В ходе ряда экспериментов было доказано, что подтасовка поисковой выдачи может повлиять на мнение и решения до 63% респондентов⁴.

Законодательное регулирование деятельности рекомендательных сервисов

В Китае с 1 марта 2022 г. вступил в силу Регламент работы рекомендательных систем. В США и ЕС существуют проекты законов и актов, направленных на регулирование работы рекомендательных сервисов. Законодательные инициативы этих стран характеризуют желание регулировать работу соответствующих алгоритмов.

² Новостное интернет-издание Lenta.ru. Новая нейросеть Facebook начиталась разговоров пользователей и стала расистской.

<https://lenta.ru/news/2023/05/18/rabotat/>

³ Принадлежит корпорации Meta, признанной экстремистской и запрещенной в России.

⁴ Новостное интернет-издание Lenta.ru. «Перепрограммировать человечество»: Google и Facebook управляют миллиардами людей. Зачем им абсолютная власть над миром?

<https://lenta.ru/articles/2021/10/14/epstein/>

Регламент работы рекомендательных систем **КНР** устанавливает, что компании, использующие рекомендательные алгоритмы, должны отправить их код на проверку регулятору КНР, а также самостоятельно проводить оценку безопасности алгоритмов согласно установленному государством регламенту. Алгоритмические рекомендательные сервисы, предлагающие новостную информацию, должны получить специальную лицензию.

На операторов рекомендательных сервисов с возможностью выражения общественного мнения или с возможностями социальной мобилизации возлагается обязанность в течение 10 рабочих дней с момента предоставления услуг сообщить в Центральную комиссию по кибербезопасности и информатизации название оператора, форму услуги, область применения, тип алгоритма, отчет об оценке безопасности, контент, предназначенный для публикации, и другую необходимую информацию.

Запрещена реклама фейковых новостей, провокация на совершение противоправных действий, нарушение общественного порядка, подрыв национальной безопасности. Информация, предоставляемая рекомендательными сервисами несовершеннолетним, не должна провоцировать на имитацию опасного поведения или действий, нарушающих общественную мораль, привести к вредным наклонностям или иным образом повлиять на физическое и психическое здоровье.

Пользователи должны иметь возможность:

- отказаться от получения рекомендаций, сформированных с помощью алгоритмов;
- выбрать или удалить метки, которые используют алгоритмы для формирования рекомендаций.

Компании обязаны предоставить возможность безопасного использования рекомендательного сервиса для пожилых людей, защитив их от мошенничества.

В Европейском союзе в ноябре 2022 г. вступил в силу Закон о цифровых услугах — Digital Services Act (DSA)⁵, который обязывает онлайн-платформы предоставлять по запросу:

- ♦ данные, необходимые для оценки рисков и возможного вреда, наносимого системами платформы;
- ♦ данные о точности, функционировании или тестировании алгоритмических систем для модерации контента, рекомендательных и рекламных сервисов;
- ♦ данные о процессах и выводах систем модерации контента или внутренних систем по обработке жалоб пользователей.

Онлайн-платформы как минимум один раз в год должны проводить оценку рисков, связанных с работой или развитием их сервисов в ЕС.

⁵ Сайт Европейской комиссии. Proposal for a Regulation on a Single Market For Digital Services (Digital Services Act) https://ec.europa.eu/info/sites/default/files/proposal_for_a_regulation_on_a_single_market_for_digital_services.pdfM



Кроме того, онлайн-платформы должны предоставлять пользователям доступную для понимания информацию об основных параметрах и критериях, используемых в рекомендательных системах, и предоставлять по крайней мере один вариант работы системы, не основанный на профилировании пользователей.

Онлайн-платформы обязаны по запросу обеспечивать своевременный доступ к любым системам, подпадающим под регулирование DSA.

В США на рассмотрении находится законопроект (последние слушания проходили в конце 2022 г.) о запрете использования большей части персональных данных (например, о расе, поле, религиозных убеждениях и других персональных данных из баз брокеров данных) для проведения таргетированных рекламных кампаний в интернете. Определять целевую аудиторию для рекламодателей предлагается на основе «общих данных о местоположении на уровне города или штата», а также «на основе контента, с которым взаимодействует пользователь».

Согласно другому законопроекту, призванному защитить пользователей от манипулирования, онлайн-платформы могут обязать технически обеспечивать возможность использования версии платформ без рекомендательных алгоритмов. Пользователь должен выразить однозначное согласие на использование своих данных в целях работы рекомендательных алгоритмов. Документ очерчивает круг данных, которые не могут быть использованы в этих целях (например, данные истории привязанного устройства).

В России депутаты Госдумы работают над законопроектом о рекомендательных сервисах, не позволяющим манипулировать потребителями, создавать аномальный интерес к товарам и влиять на общественное мнение.

Законодательное регулирование рекомендательных сервисов необходимо для защиты потребителей от непрозрачных алгоритмов отбора контента, способных исказить реальное положение дел и приводить к пугающим последствиям, вплоть до реального нанесения вреда здоровью и психике человека.

Ключевой аспект регулирования рекомендательных систем в международном законодательстве — вопрос использования персональных данных пользователей.

Пользователь должен быть в доступной форме и в полной мере проинформирован о работе рекомендательного алгоритма, дать согласие на обработку персональных данных и иметь возможность его отозвать.

Все данные, приведенные в настоящем документе, взяты из электронных материалов СМИ, официальной информации органов исполнительной власти и публикаций организаций и предприятий.

Научно-технический центр ФГУП «ГРЧЦ»
Блок развития

ОПРЕДЕЛЕНИЕ РЕКОМЕНДАТЕЛЬНЫХ ТЕХНОЛОГИЙ И СЕРВИСОВ

Согласно дорожной карте «Новые коммуникационные интернет-технологии», подготовленной «Ростелекомом» в рамках соглашения с Правительством РФ и утвержденной Правительственной комиссией по цифровому развитию, к поисково-рекомендательным технологиям относятся:

- ♦ сервисы, предлагающие персонализированный контент для развития личности потребителя на базе рекомендательных технологий, в том числе сервисы для индивидуального прогнозирования карьерного развития и динамического мониторинга состояний (настроения человека);
- ♦ рекомендательные сервисы, основанные на программном комплексе коллаборативной фильтрации (прогнозы поведения пользователей, исходя из накопленной информации об интересах и вкусах других пользователей);
- ♦ программный комплекс управления процессами извлечения, преобразования и загрузки данных для подключения к рекомендательным сервисам конечных потребителей.

Также на основе поисково-рекомендательных технологий создаются:

- ♦ сервисы интеллектуального поиска по различным видам медиаконтента, в том числе интеллектуального анализа видеопотока, системы выявления цепочек распространения инфоповодов и идентификации инфоповодов, распространяемых ботами;
- ♦ нейронные сети, позволяющие в автоматическом режиме проводить оценку профессиональных качеств и компетенций кандидатов на вакансии на основе открытых резюме, выбирать оптимальное соотношение «соискатель-вакансия» и формировать (при необходимости) рекомендации соискателям для достижения соответствия требованиям работодателя;
- ♦ сервисы по формированию тематических сообществ (подбор собеседников по интересам и потребностям) в социальных сетях и других коммуникационных сервисах, например, в области науки, образования, профессиональной деятельности, волонтерства, творчества, спорта.

Согласно законодательству КНР, к технологиям алгоритмических рекомендаций относятся типы технологий и алгоритмов, которые осуществляют генерацию и синтез, индивидуализированный подбор, уточнение последовательности вывода, фильтрацию поиска и т. д. для предоставления пользователям информации.

В законопроекте *Filter Bubble Transparency Act*, представленном на рассмотрение Сената США, введено понятие «алгоритмическая система ранжирования». Согласно документу, это вычислительный процесс, в том числе основанный на алгоритмическом принятии решений, машинном обучении, статистическом анализе или других методах обработки данных или искусственного интеллекта, используемый для определения порядка и способа предоставления информации пользователю на закрытой интернет-платформе, включая ранжирование результатов поиска, предоставление рекомендаций по контенту, отображение сообщений в социальных сетях или любой другой метод автоматического выбора контента.

Согласно предлагаемым изменениям в Digital Services Act (ЕС), суть рекомендательных систем состоит в алгоритмическом отображении, ранжировании и приоритизации информации для получателя, а также различении текстовой и визуальной информации или иной обработке информации, предоставленной получателями.

РЫНОК РЕКОМЕНДАТЕЛЬНЫХ ТЕХНОЛОГИЙ

Мировой рынок рекомендательных технологий в 2020 г. был оценен в 1,77 млрд долл. США, прогнозируемые темпы роста на период с 2021 по 2028 гг. составляют около 33% ежегодно. К 2028 г. объем рынка должен достигнуть значения в 17,3 млрд долл. США⁶. Наиболее крупным рынком для рекомендательных систем является регион Северной Америки, его доля от общемирового составляет 33%. Объем российского рынка в 2020 г. оценивался в 6 млрд руб., к 2024 г. прогнозируется его увеличение до 21,7 млрд руб.⁷

Рекомендательные сервисы на основе коллаборативной фильтрации составляют около 40% рынка за счет их преобладающего использования в системах электронной коммерции. Эту систему фильтрации использует, например, интернет-сервис потокового аудио Spotify.

В период 2021–2028 гг. прогнозируется ежегодный рост на 34,4% гибридных рекомендательных систем, используемых, например, Netflix.

ТИПЫ РЕКОМЕНДАТЕЛЬНЫХ СЕРВИСОВ И ИСПОЛЬЗУЕМЫЕ МЕТОДЫ

В процессе работы рекомендательные системы собирают данные о пользователях, используя сочетание явных и неявных методов. Примеры явного сбора данных:

- ◆ запрос у пользователя оценки объекта по дифференцированной шкале;
- ◆ запрос у пользователя ранжировки группы объектов от наилучшего к наихудшему;
- ◆ предъявление пользователю двух объектов с вопросом о том, какой из них лучше;
- ◆ предложение создать список любимых объектов пользователя.

Примеры неявного сбора данных:

- ◆ наблюдение за тем, что просматривает пользователь в интернет-магазинах или базах данных другого типа;
- ◆ ведение записей о поведении пользователя онлайн;
- ◆ отслеживание содержимого компьютера пользователя.

⁶ Здесь и далее сумма в рублях рассчитывалась по курсу ЦБ на 12.01.2022, где \$1 был равен 74,8355 руб.

⁷ Сетевое издание «Синьютс». Как в России будут поддерживаться проекты по созданию и обнаружению DeepFake. https://www.cnews.ru/articles/2021-12-14_kak_v_rossii_budut_podderzhivatsya

Фильтрация на основе контента

Рекомендательные алгоритмы на основе контента рассматривают рекомендацию как проблему классификации для конкретного пользователя, изучают классификатор симпатий и антипатий пользователя, указанных в профиле, и делают вывод на основе характеристик элемента, существующих в рекомендательной системе.

Для создания профиля пользователя система в основном ориентируется на два типа информации:

- ◆ модель предпочтений пользователя;
- ◆ история взаимодействия пользователя с рекомендательной системой.

Таким образом, эти алгоритмы пытаются порекомендовать товары или контент, похожие на те, которые пользователю нравились в прошлом или которые он изучает в настоящее время.

Один из основных плюсов данной системы — относительно небольшой рабочий объем данных, независимость от данных пользователя и легкость в их интерпретации.

Коллаборативная фильтрация

Принцип коллаборативной фильтрации основан на идее о том, что пользователи, имевшие схожие вкусы в прошлом, будут иметь их и в дальнейшем. Система генерирует рекомендации на основании рейтинговых профилей для разных пользователей или элементов, т. е. находит пользователей или элементы с историей оценок, аналогичной текущему пользователю или элементу, и генерирует рекомендации на основе этой схожести.

Ключевое преимущество коллаборативной фильтрации в том, что данный подход не полагается на анализируемый компьютером контент и способен с достаточной точностью рекомендовать сложные элементы, такие как, например, фильмы, без понимания самого элемента.

Основные проблемы данной модели — «холодный старт» нового продукта, когда требуются первичные отзывы и характеристики для начала его использования в рекомендациях, а также потребность в больших вычислительных мощностях для проведения огромного количества расчетов при составлении рекомендаций.

Сессионные системы

Такие системы составляют рекомендации только на основании действий и истории просмотров пользователя в рамках текущей сессии, их применяют, когда доступ к истории (просмотры, клики, покупки) невозможен либо не имеет значения.

Сервисы на основе множества критериев

Вместо того, чтобы создавать рекомендации на основе имеющихся данных, сервисы пытаются предсказывать другие данные о пользователе, используя информацию о предпочтениях по нескольким критериям.

Сервисы на основе машинного обучения с учителем

Ответные действия пользователя служат подкреплением для обучения системы. Например, клик по рекомендации системы, приобретение продукта по рекомендации системы.

Гибридные системы

Большинство рекомендательных сервисов используют гибридный подход, объединяя коллаборативную, контентную фильтрацию и другие методы. Варианты гибридизации: использование разных подходов последовательно с последующим объединением, добавление одного подхода в другой на определенном этапе, объединение нескольких подходов на этапе построения модели.

Техники гибридизации⁸:

Взвешенный подход	Веса внутри модели, присвоенные определенным параметрам или элементам, постепенно корректируются в соответствии со степенью, в которой оценка пользователем совпадает с оценкой, предсказанной рекомендательной системой.
Коммутационный подход	Метод изменения используемой рекомендательной модели в зависимости от эвристики, отражающей способность рекомендателя дать хороший прогноз рейтинга.
Смешанный подход	Использование нескольких видов фильтрации для составления рекомендации, например, фильтрация на основе контента при существовании проблемы «холодного старта» или недостатка информации дополняется использованием исторических данных о пользователе.
Комбинация признаков	Признаки, полученные из разных источников знаний, объединяются вместе и передаются в единый алгоритм рекомендаций.
Дополнение признаков	Определяется признак или набор признаков, которые затем становятся частью входных данных для следующего метода.
Каскад	Рекомендации одного метода уточняются другим методом рекомендаций.
Метауровень	Применяется один метод рекомендаций и создается модель, которая затем становится входными данными, используемыми для следующего метода.

⁸ A Survey of Recommendation Systems: Recommendation Models, Techniques, and Application Fields
<https://www.mdpi.com/2079-9292/11/1/141/html>

Согласно ряду исследований, именно гибридные системы дают наиболее точный результат. Для создания рекомендательных сервисов используются следующие модели и алгоритмы⁹:

- ◇ классическая регрессионная модель при использовании машинного обучения с учителем;
- ◇ использование разреженной матрицы;
- ◇ метод k-ближайших соседей;
- ◇ модели на основе правил;
- ◇ метод чередующихся наименьших квадратов;
- ◇ метод опорных векторов;
- ◇ рекуррентные нейронные сети;
- ◇ глубокие нейронные сети;
- ◇ и другие.

Оценку рекомендательных систем редко проводят на фиксированном наборе данных, поскольку невозможно точно предсказать реакцию реальных пользователей. Для измерения эффективности рекомендательных систем и сравнения различных подходов доступны три типа оценок: исследования пользователей, онлайн-оценки (A/B-тесты) и офлайн-оценки.

Пользовательские исследования имеют небольшой масштаб. Нескольким десяткам или сотням пользователей представляются рекомендации, созданные с помощью различных подходов, а затем пользователи определяют лучшие рекомендации.

В A/B-тестах рекомендации обычно показываются тысячам пользователей реального продукта. Система случайным образом выбирает как минимум два разных подхода для создания рекомендаций. Эффективность измеряют неявными показателями эффективности, такими как коэффициент конверсии или рейтинг кликов.

Офлайн-оценки основаны на исторических данных, т. е. наборе данных, который содержит информацию о том, как пользователи ранее оценивали, например, фильмы.

Кроме точности, для оценки моделей важны разнообразие предлагаемых рекомендаций, возможность повторной оценки со стороны пользователя, защита персональных данных, новизна рекомендаций (система предлагает неочевидные элементы), степень доверия пользователя, возможность оценки предложенной рекомендации (участие пользователя в оценке системы).

⁹ Более подробная информация о модели и алгоритмах для построения моделей ИИ представлена в материале НТЦ «Объяснимый ИИ»
<https://rdc.grfc.ru/2020/12/explainable-ai/>

РЕКОМЕНДАТЕЛЬНЫЕ АЛГОРИТМЫ В СОЦИАЛЬНЫХ СЕТЯХ

В основе рекомендательных алгоритмов лежит использование искусственного интеллекта, будь то простые модели логистической регрессии или наиболее продвинутые версии глубоких нейронных сетей. Поэтому можно смело утверждать, что ИИ — ключевой компонент популярных социальных сетей.

Facebook использует ИИ, чтобы понимать текст на разных языках, распознавать и классифицировать изображения, а также размещать контент и рекламу в ленте пользователей. Система искусственного интеллекта DeepFace с точностью более 97% определяет изображения и лица и используется в визуальном контенте. Система DeepText AI теперь может понимать более 20 языков.

Компания полагается на ИИ для анализа предпочтений и поведения пользователей и прогнозирования интереса пользователя к контенту и рекламе.

В 2010 г. для формирования релевантной для пользователя новостной ленты использовалось всего три параметра, а уже через несколько лет алгоритм стал использовать около ста тысяч дополнительных параметров. С тех пор совершенствование алгоритмов не прекращалось.

Теперь в расчет идут все действия пользователя: сколько секунд он просматривал пост, куда нажимал, на каких словах заострял внимание. Также учитывается модель смартфона, технические характеристики устройства, с которого человек выходит в соцсети, а при слабом сигнале Wi-Fi пользователю покажут больше текстовой информации, потому что она быстрее прогрузится.

Полный список параметров найти невозможно, поскольку соцсеть его тщательно скрывает.

Компания делает ставку на извлечение прибыли из технологий искусственного интеллекта. Чем больше пользователь взаимодействует с контентом, тем дольше он остается на платформе. Это позволяет Facebook показывать больше рекламы и собирать обширные данные о предпочтениях пользователей.

Instagram¹⁰ использует машинное обучение для идентификации визуальных элементов и таргетинга. Нейросеть анализирует взаимодействие пользователя с платформой, а затем прогнозирует, с какой следующей частью контента он будет взаимодействовать больше всего. Искусственный интеллект применяют и для пометки и блокировки спам-ботов, собирающих лайки и клики.

YouTube использует нейросети, чтобы рекомендовать видео, модерировать проблемный контент и поддерживать процесс создания контента, включая машинную транскрипцию и видеоэффекты.

Искусственный интеллект может пометать и удалять нежелательный контент практически в режиме реального времени. Это позволяет автоматически удалять вредоносный контент до того, как его увидит множество людей.

TikTok использует подход «сначала ИИ». Каждый фрагмент контента, который пользователь видит на платформе, рекомендован искусственным интеллектом.

¹⁰ Принадлежит корпорации Meta, признана экстремистской и запрещена в РФ.

Нейросеть обучается при каждом взаимодействии пользователя с контентом. Благодаря большому числу пользователей она быстро собирает огромный массив поведенческих данных.

Twitter с применением ИИ персонализирует ленту и борется с расистскими, гомофобными, исламофобскими и другими неуместными высказываниями, для чего использует IBM Watson¹¹ и обработку естественного языка (NLP). IBM Watson может анализировать миллионы непристойных и неуместных сообщений за секунды.

LinkedIn использует машинное обучение, чтобы предлагать рекомендации по работе, людям, с которыми пользователь мог бы связаться, и показывать персонифицированную ленту новостей.

Snapchat использует компьютерное зрение и технологии ИИ, чтобы отслеживать черты лица и накладывать фильтры, которые перемещаются вместе с лицом пользователя в режиме реального времени.

ВКонтакте использует рекомендательные сервисы для формирования ленты пользователя, но функцию можно отключить и просматривать новости в хронологическом порядке.

Лента **Дзен** также выстраивается согласно предполагаемым интересам пользователя.

Практически любая онлайн-платформа дает маркетологам возможность показывать платную рекламу на основе точного демографического и поведенческого таргетинга. С помощью ИИ объявления оптимизируются для повышения конверсии.

Некоторые инструменты ИИ могут даже предсказать, какая реклама будет эффективнее, прежде чем она будет запущена в социальных сетях.

ИИ может управлять рекламными кампаниями, корректируя таргетинг, креативы и бюджет в режиме реального времени для оптимизации производительности.

Однако при распространении контента ИИ часто отдает предпочтение спорной, разжигающей вражду и недостоверной информации, поскольку именно такой контент привлекает больше внимания и ведет к повышению вовлеченности во взаимодействие с платформой.

Исследователи неоднократно отмечали эту закономерность в отношении работы Facebook AI¹². Алгоритм, нацеленный на распространение контента с максимальной вовлеченностью, в итоге продвигал среди пользователей фейковые новости и риторику ненависти. Эту проблему, по мнению экспертов, может решить ручная модерация контента.

Другим следствием использования рекомендательных алгоритмов может быть создание «эхокамер» — окружение пользователя информацией схожего характера. Это происходит, когда рекомендательная система предлагает контент других пользователей на основании схожести взглядов и интересов. Человек оказывается огражден от информации другого типа или от критической информации в отношении собственных взглядов. В результате социальные сети могут спровоцировать политическую поляризацию и продвигать дезинформацию.

¹¹ Одна из первых когнитивных систем в мире.

<https://habr.com/ru/companies/ibm/articles/266015/>

¹² Принадлежит корпорации Meta, признанной экстремистской и запрещенной в России.

Предложения возможных «друзей» и «подписок», основанные на структурной схожести профилей (например, большое количество общих друзей), приводят к поляризации мнений и повышению коэффициента кластеризации онлайн-платформы (группы пользователей становятся все более изолированными)¹³.

Сотрудница Facebook, социолог Моника Ли отмечала, что 64% всех новых подписчиков экстремистских групп пришли из рекомендаций¹⁴.

Доктор наук Гарвардского университета, ведущий психолог-исследователь в Американском институте поведенческих исследований и технологий в Калифорнии Роберт Эпштейн после проведения первого эксперимента¹⁵ с результатами поисковой выдачи выяснил, что ее подтасовка может повлиять на решение почти 48% неопределившихся респондентов на выборах. Второй эксперимент показал, что таких людей может быть больше — до 63%. При этом предвзятость поисковой выдачи оказывается довольно легко скрыть, поскольку индивидуально подобранные для пользователя результаты (т. н. «эфемерный опыт») нигде не хранятся.

Команда Эпштейна перед выборами 2020 г. в США разработала систему мониторинга поисковой выдачи. В ходе предвыборной гонки исследователи отметили существенный перекося в результатах поиска Google в пользу демократов.

ЗАКОНОДАТЕЛЬСТВО В ОТНОШЕНИИ РЕКОМЕНДАТЕЛЬНЫХ СЕРВИСОВ

КИТАЙ

В августе 2021 г. Управление по вопросам киберпространства КНР опубликовало проект Регламента работы рекомендательных систем. Документ вступил в силу 1 марта 2022 г.

Согласно документу, к рекомендательным относятся системы, использующие алгоритмы генеративного или синтетического типа для предоставления информации пользователям, формирования персонализированных рекомендаций, ранжирования или принятия решений, фильтрации поиска.

Операторы рекомендательных сервисов с возможностью выражения общественного мнения или социальной мобилизации должны в течение 10 рабочих дней с момента предоставления услуг сообщить название оператора, форму услуги, область применения, тип

¹³ Link recommendation algorithms and dynamics of polarization in online social networks
<https://www.pnas.org/doi/10.1073/pnas.2102141118>

¹⁴ Новостное интернет-издание Lenta.ru. «Мы все — подопытные крысы» Соцсети решают, что читают и смотрят миллиарды людей. Чем это опасно для человечества?
https://lenta.ru/articles/2021/10/11/algorithm_1/

¹⁵ Новостное интернет-издание Lenta.ru. «Перепрограммировать человечество» : Google и Facebook управляют миллиардами людей. Зачем им абсолютная власть над миром?
<https://lenta.ru/articles/2021/10/14/epstein/>

¹⁶ Сайт Digichina. Translation: Internet Information Service Algorithmic Recommendation Management Provisions – Effective March 1, 2022
<https://digichina.stanford.edu/work/translation-internet-information-service-algorithmic-recommendation-management-provisions-effective-march-1-2022/>

алгоритма, отчет об оценке безопасности алгоритма (проводится согласно установленному государством регламенту), контент, предназначенный для публикации, и другую подобную информацию.

Обработка поданной оператором информации проводится в течение 30 рабочих дней национальными и провинциальными, автономными районными и муниципальными управлениями кибербезопасности и информатизации.

Среди новых правил регулирования рекомендательных сервисов:

- ◆ Компании не должны использовать рекомендательные алгоритмы с противозаконной целью, недопустимы манипуляции сознанием пользователей с целью провоцирования их на совершение противоправных действий, нарушение общественного порядка, подрыва национальной безопасности.
- ◆ Алгоритмические рекомендательные сервисы, предлагающие новостную информацию, должны получить специальную лицензию, рекомендации фейковых новостей запрещены.
- ◆ Поставщики услуг алгоритмических рекомендаций должны усилить управление информационной безопасностью; создать и заполнить базы данных признаков, которые будут использоваться для выявления незаконной и вредоносной информации; сформировать стандарты и нормы для заполнения баз данных.
- ◆ Компании обязаны информировать пользователей о принципах, цели и порядке работы рекомендательного сервиса.
- ◆ Пользователи должны иметь возможность отказаться от получения рекомендаций, сформированных с помощью алгоритмов.
- ◆ Пользователи должны иметь возможность выбрать или удалить метки, используемые алгоритмом для формирования рекомендации.
- ◆ Операторы рекомендательных систем должны создать сервисы для подачи жалоб, установить сроки их обработки и обратной связи.
- ◆ Компании обязаны защитить пожилых людей от мошенничества, предоставляя возможность безопасного использования рекомендательных сервисов.
- ◆ Информация, предоставляемая рекомендательными сервисами несовершеннолетним, не должна провоцировать на имитацию опасного поведения, действий, нарушающих общественную мораль, привести к вредным наклонностям или иным образом повлиять на физическое и психическое здоровье несовершеннолетнего.
- ◆ Все компании, использующие рекомендательные алгоритмы, должны отправить их код на проверку регулятору КНР.

За нарушение правил для компаний предусмотрены штрафы от 10 до 100 тыс. юаней, возможна блокировка обновлений информации операторов рекомендательных сервисов.

ЕВРОПЕЙСКИЙ СОЮЗ

В ЕС подписан Закон о цифровых услугах (Digital Services Act, DSA). В фокусе регулятора — рекомендательные системы, используемые крупными онлайн-платформами, такими как Facebook и Google.

◆ Крупнейшие онлайн-платформы должны предоставлять пользователям:

- ◆ доступную для понимания информацию об основных параметрах и критериях, используемых в рекомендательных системах;
- ◆ возможность менять эти параметры;
- ◆ по крайней мере один вариант работы системы, не основанный на профилировании пользователей.

◆ Онлайн-платформы должны обеспечить пользователю возможность в любой момент без затруднений изменить свой выбор в отношении работы рекомендательных сервисов.

◆ Платформы должны самостоятельно регулировать и останавливать распространение с помощью рекомендательных сервисов противозаконного контента. При необходимости рекомендуется прибегать к помощи доверенных компаний для поиска противозаконной информации.

◆ В целях надзора за работой онлайн-платформ в рамках DSA специальный координатор по цифровым сервисам¹⁸ или Комиссия могут потребовать у платформы предоставлять:

- ◆ данные, необходимые для оценки рисков и возможного вреда, наносимого системами платформы;
- ◆ данные о точности, функционировании или тестировании алгоритмических систем для модерации контента, рекомендательных и рекламных сервисов;
- ◆ данные о процессах и выводах систем модерации контента или внутренних систем по обработке жалоб пользователей.

¹⁸ Координатор по цифровым сервисам в стране работы оператора или его представителя в рамках ЕС.

- ◆ Онлайн-платформы обязаны иметь в штате специалиста по комплаенсу (адаптации к закону). Ему должен быть обеспечен своевременный доступ к любым системам, подпадающим под регулирование DSA.

За нарушение норм DSA Комиссия может налагать на онлайн-платформу штрафы, в сумме не превышающие 6% от ее общего оборота в предыдущем финансовом году. При определении размера штрафа Комиссия должна учитывать характер, тяжесть, продолжительность и повторяемость нарушений.

США

В США законодатели рассматривают законопроект (последние слушания проходили в конце 2022 г.)¹⁹ против манипулирования пользователями со стороны онлайн-платформ на основании данных о них. Условное название — Filter Bubble Transparency Act.

- ◆ Платформы должны уведомлять пользователя при первом контакте с «непрозрачным» алгоритмом о методах и целях его использования, предоставлять возможность отказа от использования таких алгоритмов.
- ◆ В течение года после опубликования закона онлайн-платформы обязаны будут технически обеспечить возможность использования версии платформ без рекомендательных алгоритмов. Пользователям должна быть доступна возможность переключения между версиями платформ с и без использования рекомендательных алгоритмов.
- ◆ Пользователь должен выразить однозначное согласие на обработку персональных данных в целях работы рекомендательных алгоритмов. Документ очерчивает круг данных, которые не могут быть использованы в этих целях (например, данные истории привязанного устройства).

¹⁹ Filter Bubble Transparency Act

<https://www.congress.gov/bill/117th-congress/senate-bill/2024/text>