

Е.С. Ганюшкина

ВЛИЯНИЕ АЛГОРИТМОВ ФИЛЬТРАЦИИ СОЦИАЛЬНЫХ СЕТЕЙ НА ФОРМИРОВАНИЕ ОБЩЕСТВЕННОГО МНЕНИЯ

E. Ganyushkina. Influence of filtering algorithms of social networks on the formation of public opinion

Аннотация. Статья посвящена анализу потенциального влияния алгоритмов фильтрации социальных сетей на формирование общественного мнения. В ней рассмотрены используемые типы алгоритмических механизмов, принцип работы, а также их способность противостоять дезинформации. Приведены примеры использования алгоритмов фильтрации на популярных социальных площадках – ВКонтакте, YouTube, Яндекс.Дзен. А также представлены возможные способы улучшения эффективности и качества их работы в будущем.

Ключевые слова: общественное мнение, социальные сети, алгоритмы фильтрации, персонализация информации, дезинформация.

Контактные данные: Россия, 191023, Санкт-Петербург, Садовая ул., д.21;
e-mail: katerina131299@gmail.com

Abstract. The article is devoted to the analysis of the potential influence of social network filtering algorithms on the formation of public opinion. It examines the types of algorithmic mechanisms used, how they work, and their ability to counter disinformation. Examples are given of the use of filtering algorithms on popular social platforms - VKontakte, YouTube, Yandex.Zen. Possible ways to improve the efficiency and quality of their work in the future are also presented.

Keywords: public opinion, social networks, filtering algorithms, personalization of information, disinformation

Contact information: Russia, 191023, St. Petersburg, Sadovaya st., 21;
e-mail: katerina131299@gmail.com

С появлением интернета началась новая эпоха. Люди стали проводить в социальных сетях огромное количество времени, а для современного поколения детей, подростков даже нет разделения на виртуальную и реальную жизнь, в первой они чувствуют себя даже комфортнее. Интернет дает много преимуществ, например, доступ к информации любого вида – стоит только захотеть и можно посмотреть концерт, проходящий на другом конце света, или оказаться на премии «Оскар». Но информации становится все больше и больше. Здесь на помощь приходят разработанные алгоритмы фильтрации данных, которые по задумке должны ограничивать бесконечный информационный поток. Однако существуют опасения использования этих фильтрационных алгоритмов, заключающиеся в определении их влияния на человека и на общество в целом.

Роль социальных сетей в формировании общественного мнения

Роль социальных сетей Интернета в жизни общества не стоит преувеличивать или преуменьшать. Социальные сети – часть современной информационной реальности. Они могут быть использованы, как и любой другой инструмент коммуникации, для разных целей.

Сегодня социальные сети служат платформой для формирования общественного мнения и для его выражения. Большую часть знаний о мире, в том числе о политике и социально-экономических процессах, люди получают из социальных сетей, также с помощью них они могут быстро делиться своими мыслями с другими.

Ганюшкина Екатерина Сергеевна – магистрант, Санкт-Петербургский государственный экономический университет

E. Ganyushkina - master's student of St. Petersburg State Economic University

© Ганюшкина Е.С., 2024

Основной особенностью формирования общественного мнения с помощью социальных сетей является возросшая скорость распространения информации, люди могут за считанные секунды связаться друг с другом и поделиться новостью. Оставить комментарий под каким-либо постом – дело, занимаемое меньше минуты. Например, о том, что Twitter заблокировал аккаунт Д. Трампа, миллионы людей узнали спустя долю секунды.

Также еще одной отличительной чертой стало то, что с помощью социальных сетей множество людей может участвовать в дискуссии по какому-либо волнующему обществу вопросу, высказывать свое мнение и в процессе формировать общественное. Удивителен тот факт, что социальные сети объединяют весь мир, а простота и удобство их использования говорит о том, что в будущем они никуда не исчезнут.

Однако есть определенные риски, связанные с формированием общественного мнения с помощью социальных сетей.

Во-первых, социальные сети могут способствовать распространению дезинформации и фейковых новостей, что может вводить людей в заблуждение и способствовать формированию неправильного общественного мнения.

Во-вторых, это алгоритмы фильтрации контента, на которых построены умные ленты, предлагающие новости пользователям. Возникает ряд важных вопросов: насколько мнение людей формируется содержанием того, что показывается в социальных сетях и какого влияние этих алгоритмов на формирование персонализированного общественного мнения, под которым понимается уникальное явление, которое заключается в том, что отдельные люди склонны формировать свои суждения о какой-либо проблеме на основе мнения, которое исходит от общества в целом.

Краткий обзор влияния алгоритмов на информационный поток

Фильтрация в социальных сетях означает, что контент, предлагаемый пользователям, подвергается цензуре или распределяется с помощью алгоритма с целью соответствовать тому, что предпочитает отдельный пользователь.

Влияние алгоритмов фильтрации на общественное мнение все больше волнует современных ученых. Например, Эли Парайзер исследовал влияние алгоритмов персонализации контента на формирование общественного мнения. В своей книге «The Filter Bubble: How the New Personalized Web is Changing What We Read and How We Think» автор исследует феномен фильтров-пузырей. А в своем выступлении для TED [12] он приводит пример опасности фильтрационных алгоритмов. Исследователь попросил друзей посмотреть, что выдаст их Google по запросу «Египет», и прислать результаты ему. Получились совершенно разные результаты. У первого друга не было даже намеков на происходящие в Египте протесты, а у второго была информация только про них. А ведь эта была одна из главных новостных повесток.

Также проводились исследования по этой теме крупными университетами, например, в университете Беркли занимались выявлением и уменьшением алгоритмического биаса (предвзятости), который влияет на общественное мнение. Согласно опубликованной работе под названием «Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms» [2, с. 1], алгоритмы используют объемы макро- и микроданных, чтобы влиять на решения, затрагивающие людей, в решении целого ряда задач, от составления рекомендаций по фильмам до оказания помощи банкам в определении кредитоспособности физических лиц. Авторы статьи говорят о том, что принятые алгоритмами решения могут не совпадать с ожиданиями людей, и даже могут стать причиной усиления предубеждений в обществе, приводя в пример используемые судьями в США автоматизированные оценки рисков для вынесения приговоров, которые могут приводить к тому, что людям с кожей отличающегося цвета может быть увеличена сумма залога или срок тюремного заключения. А Pew Research Center [1] проводил исследование влияния алгоритмов Facebook (запрещенная в РФ соцсеть) на просмотр новостей. Специалисты привели данные опроса среди

американских граждан, которые показывают, что большинство людей не знают, как работает умная лента в социальной сети, и почему они видят именно ту информацию, которую им показывают.

Алгоритмы фильтрации в социальных сетях – это тип программного обеспечения, которое использует определенные критерии для фильтрации и поиска конкретной информации из большого объема данных в интернете.

Они могут оказывать влияние на информационный поток, например, ограничивая доступ к разным типам информации. Это в свою очередь может привести к неприятным последствиям.

– Ограничение информационного пространства пользователя. Другими словами, если какой-то контент попадет в раздел «нежелательно», то вся похожая информация не будет показана пользователю. Также это приведет к тому, что снизится количество разнообразных точек зрения по какому-либо определенному вопросу.

– Дезинформирование. Алгоритмы основаны на машинном обучении, которое построено на исторических данных, которые могут содержать стереотипы, либо просто устаревать. Также алгоритмы не всегда могут распознать ложную информацию.

– Влияние на свободу слова посредством ограничения доступа к ней.

– Манипулирование информационными потоками.

Работа алгоритмов

Существует несколько видов алгоритмов фильтрации в социальных сетях.

1. Алгоритмы, использующие списки нежелательных или запрещенных элементов (например, запрещенные слова), которые при выявлении либо блокируются, либо вообще удаляются.

2. Алгоритмы машинного обучения. Они используют статистические методы и анализ данных для определения нежелательных элементов. Машинное обучение – область искусственного интеллекта, которая позволяет компьютерам учиться на основе данных и опыта, без сложного программирования. Сегодня машинное обучение используют практически в каждом приложении, например, для рекомендательных систем.

3. Алгоритмы семантической фильтрации. Они анализируют содержание информации и соответствует ли оно заданным критериям. Яркий пример, анализ по определенным фразам.

4. Алгоритмы нечеткой логики. Используются для выявления соответствия заданным критериям.

Все алгоритмы фильтрации работают по-разному, но общие процессы также имеются. Все начинается со сбора информации о пользователе. Обычно используются данные местоположения, возраста, интересов, друзей и т.д. Далее на основе собранной информации создается модель поведения пользователя, которая учитывает его интересы и другие характеристики. На основе этой модели идет фильтрация контента, который показывается приложением, определяется, что соответствует предпочтениям, а что нет. Кроме того, алгоритм ранжирует информацию, у той, которой процент соответствия больше, будет показана первой всех. Модель поведения постоянно обновляется, учитывая новые данные, что позволяет ей работать эффективнее и точнее.

Примеры в популярных платформах

Например, популярная социальная сеть ВКонтакте использует для своей умной ленты алгоритмы машинного обучения. Как работают эти алгоритмы можно узнать из опубликованных «Правил применения рекомендательных технологий социальной сети ВКонтакте» [8]. Из них ясно, что алгоритмы применяются для того, чтобы предоставлять людям качественную и релевантную информацию. В основном собираются сведения о предпочтениях пользователей из Российской Федерации, которые определяются по их действиям в социальной сети, например, какой контент они читают чаще других. Алгоритмы

учитывают отметки «мне нравится», «поделиться», «не интересно», частоту просмотра той или иной информации, комментарии, вступления в сообщества или подписки и т.д. Кроме того, используется данные аккаунта, например, местоположение, чтобы рекомендовать контент, релевантный месту жительства. Важно то, что для продвижения публикации играют роль обсуждения, посты с большим количеством комментариев будут продвигаться выше по ленте и их увидит большее количество людей. Особенностью ВК является возможность отключить умную ленту в настройках.

Похожие алгоритмы можно наблюдать у YouTube. Эта социальная сеть полностью ориентирована на рекомендации. На главной странице высвечиваются персонализированные рекомендации, то, что соответствует интересам пользователя, или популярно у людей со схожими предпочтениями. Для этих рекомендаций важно, насколько видео нравится людям со схожими интересами и как часто пользователь смотрит видео на данную тему. Они учитывают также количество просмотров, насколько быстро растет эта цифра, время публикации видео, сравнение с другими видео на канале.

Интересные алгоритмы использует «Яндекс.Дзен»: фильтрация содержимого (проверяются характеристики и содержание публикации) и коллаборативная фильтрация (показ релевантного контента, похожего на тот, который пользователь оценил ранее). Особенностью платформы является то, что авторы здесь должны публиковать контент, а алгоритмы обеспечат их читателями.

Влияние алгоритмов на формирование мнения пользователя

Многие исследователи интересуются проблемой влияния технологий фильтрации информации на отдельную личность и общество в целом. Много внимания уделяют феномену, который называют «пузырем фильтров». Под ним подразумевается ситуация, когда пользователь попадает в замкнутое персонализированное информационное пространство. Ему ограничивают доступ информации определенного вида. В. Грушевская причиной этого феномена считает психологические особенности людей, которые хотят видеть только ту информация, которая подтверждала бы их взгляды [6, с. 396]. Ученые также отмечают факт, что рекомендательная система выдает достаточно однородную информацию, а также обвиняют ее в предвзятости. Это связано с тем, что, например поисковые платформы могут выдавать, к примеру, одни новостные сайты чаще других.

Также много опасений вызывает манипулятивный потенциал всех алгоритмов фильтрации информации. Например, С. В. Володенков пишет, что данные пользователей собираются «для последующего их использования в интересах субъектов глобального политического управления» [5, с. 342]. Кроме того, производители приложений, сервисов и даже продавцы товаров заинтересованы в том, чтобы люди проводили больше времени в интернете для этого используются разные методы, например, вовлечение пользователей в сетевые дискуссии.

Социальные сети – один из инструментов социализации современного общества. В них могут формироваться ценностные ориентиры, традиции, даже гражданская позиция. Это достигается с помощью различных текстовых и видео-постов. Яркий пример приводит В. Грушинская в своей статье, вспоминая о публикации провокационного видео-ролика на платформе YouTube под названием «дворец для П. История самой большой взятки», который сопровождался призывами к участию в протестующих акциях и всколыхнул общественное мнение.

Роль алгоритмов в распознавании и борьбе с дезинформацией

Тем не менее, распространение некоторой информации, например, фейковой, вводящей в заблуждение, должно быть ограничено. Дезинформация – это ложная информация, которая публикует с целью изменить мнение или поведение людей, запутать или даже причинить вред отдельным личностям или группам людей. Это может быть ложь или искажение фактов,

которые в конечном счете могут сформировать неверное общественное мнение и привести к серьезным последствиям.

Алгоритмы фильтрации также направлены на борьбу с дезинформацией. Они могут автоматически не допускать до появления у пользователей подозрительных публикаций. Как правило они, анализируют большие объемы информации, делая акцент на противоречиях в собранных данных и на выявлении подозрительной активности пользователей и заданных шаблонов, которые могут свидетельствовать о дезинформации.

Сегодня социальные сети вынуждены вести активную борьбу с дезинформацией. Некоторые из них нанимают в штат дополнительных работников, чтобы они проверяли публикации на достоверность на разных языках. Также социальные сети поощряют пользователей, которые сообщают о фейковых постах в поддержку. Борьба с распространением ложной информации задача политики безопасности не только каждой отдельной социальной сети, но и государства в целом.

Возможные способы улучшения алгоритмов фильтрации для создания более разнообразного информационного потока

Как правило, работу алгоритмов можно улучшить несколькими способами.

Во-первых, предоставить алгоритмам более разнообразные данные, чтобы они лучше различали разные типы информации. Например, информацию из разных источников, с разных точек зрения, и разных областей.

Во-вторых, использовать более сложные методы машинного обучения. На сегодняшний день есть несколько популярных: обучение с учителем (есть помеченные данные, на которых искусственный интеллект учится классификации и регрессии), обучение без учителя (искусственный интеллект учится анализировать данные и выявлять скрытые закономерности), онлайн-обучение (модель получает данные в реальном времени), обучение с подкреплением (модель обучается на основе обратной связи) и другие. Для улучшения алгоритмов можно чередовать разные методы машинного обучения.

В-третьих, следить за качеством информации. Данные для обучения алгоритмов должны быть полными и достоверными. Алгоритмы должны учитывать контекст, в котором находится то или иное слово или фраза. Это поможет им лучше понимать значение информации, которую они анализируют.

В-четвертых, увеличить участие человека в работе с алгоритмами, особенно в мониторинге и проверке их работы. Поможет в этом проведение качественного и регулярного аудита. У алгоритма нет другого выбора, как быть продуманным заранее. Аудит требует проверки как входных данных, так и выходных решений, и, когда они выполняются сторонним оценщиком, они могут дать представление о поведении алгоритма. Хорошей практикой является контрольный журнал работы механизма.

Работа над улучшением алгоритмов должна проводиться совместно не только с теми, кто создает их, но и с теми, на кого они влияют.

Выводы. Социальные сети прочно и надолго вошли в современную мир. Люди не могут представить свою жизнь без них. Они удобны и просты в использовании. С помощью них можно быть в курсе новостей со всего света, а также заводить новых друзей.

Однако алгоритмы фильтрации, заложенные в основу социальных сетей, представляют неосознанную угрозу для общественного мнения, формированию которого способствуют данные медиа. Алгоритмы построены таким образом, чтобы предлагать пользователю контент, основанный на его интересах, данные о которых берутся из его профиля и истории действий. Но это может привести к тому, что человек будет ограничен в информационном потоке. Есть две причины для этого: во-первых, человек всегда будет рад той информации, которая подтверждает его взгляды, на что и ориентированы встроенные алгоритмы; во-вторых, это деятельность заинтересованных сторон, например, производители продуктов

либо индивиды, которые умышленно хотят ввести людей в заблуждение путем распространения дезинформации.

Борьба с ложными новостями также входит в задачи алгоритмов фильтрации, однако эта функция все еще находится в процессе совершенствования.

Тем не менее, отказаться от данных алгоритмов не представляется возможным, т.к. они обеспечивают хоть какую-то ротацию большого информационного потока, связанного с появлением Интернета. Поэтому лучшим выходом борьбы с отрицательным влиянием работы этих алгоритмов на формирование общественного мнения является их усовершенствование.

Литература

1. Many Facebook users don't understand how the site's news feed works / Pew Research Center URL: <https://www.pewresearch.org/short-reads/2018/09/05/many-facebook-users-dont-understand-how-the-sites-news-feed-works/> (дата обращения 24.03.2024)
2. Lee N.T., Resnick P., Barton G. Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms / Brooking. 2021 29 p.
3. Pariser E. The Filter Bubble: How the New Personalized Web is Changing What We Read and How We Think. London, 2011. 304 p.
4. Авхадеев Б.Р., Воронова Л.И., Охупкина Е.П. Разработка рекомендательной системы на основе данных из профиля социальной сети «ВКонтакте». Вестник Нижневартковского государственного университета. №3. М., 2014. 68-76 с.
5. Володенков С.В. Интернет-коммуникация в глобальном пространстве современного политического управления: навстречу цифровому обществу. Проспект, М., 2021. 416 с.
6. Грушевская В. Модель фильтрации информации в социальных медиа. Журнал исследований социальной политики. Т. 20, №3, 2022. 393 – 406 с.
7. Изменится ли постинг в ВК после раскрытия информации о работе алгоритмов и что теперь учитывать. SMMPLANNER. URL: <https://smmplanner.com/blog/izmienitsia-li-postingh-v-vk-poslie-raskrytiia-informatsii-o-rabotie-algoritmov-i-chto-tiepier-uchityvat/> (дата обращения: 21.03.2024)
8. Правила применения рекомендательных технологий социальной сети ВКонтакте URL: <https://vk.com/legal/recommendations> (дата обращения: 21.03.2024)
9. Тумбинская М.В. Системный подход к обеспечению защиты от нежелательной информации в социальных сетях. Вопросы кибербезопасности №2 (20), 2017. 30-44 с.
10. Филь М.С. Социальные технологии. Новые технологии управления миром. Московский финансово-промышленный университет «Синергия», 2016. 426 с.
11. Чжан Ч., Афанасьев Г.И. основные технологии и перспективы эволюции персонализированных рекомендательных систем. E-SCIO. Информационные технологии. №4 (67). Саранск, 2022. 309-320 с.
12. Э. Парайзер: Остерегайтесь онлайн «фильтров-пузырей» URL: https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles?subtitle=ru&language=ru (дата обращения 24.03.2024)