

УДК 004.5

DOI 10.17513/snt.39861

МЕТОД КОРРЕКТИРОВКИ РЕЛЕВАНТНОСТИ ДЛЯ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ

¹Митрохин М.А., ²Мартышкин А.И., ¹Ионова Д.Н.,
¹Федяшов М.С., ¹Горынина А.В.

ФГБОУ ВО «Пензенский государственный университет», Пенза,

e-mail: mmax83@mail.ru, ionova243@gmail.com, mark02042001@mail.ru, zmey.6453@mail.ru;

²ФГБОУ ВО «Пензенский государственный технологический университет»,
Пенза, *e-mail: mai@penzgtu.ru*

В статье рассматривается метод корректировки релевантности при составлении списка рекомендаций, учитывающий необходимость синхронизации доступа к критическому ресурсу при асинхронной работе. Проанализирована чувствительность различных рекомендательных алгоритмов на основе предпочтений пользователя к таким проблемам, как холодный старт, пропущенные значения, информационный пузырь. Группу методов подхода на основе данных представляли коллаборативная фильтрация, фильтрация на основе контента и фильтрация на основе знаний. В качестве примеров методов подхода на основе моделей рассматривались байесовские классификаторы, матричная факторизация, нейронные сети. Метод на основе байесовских классификаторов стал предпочтительным для дальнейшего рассмотрения за счет скорости реализации и упрощенного процесса проектирования. В статье описан метод решения проблемы холодного старта и информационного пузыря для байесовского рекомендательного алгоритма на основе корректирующих показателей: отрицательные веса, коэффициент затухания и коэффициент новизны, которые дают возможность разнообразить подборки пользователя. В дополнение к этому был предложен механизм синхронизации процесса изменения массива отрицательных весов, являющегося критическим ресурсом при условии совместного использования корректирующих показателей. В качестве наглядного представления механизма редактирования массива отрицательных весов был использован метод формализации с применением событийных недетерминированных автоматов.

Ключевые слова: рекомендательные системы, байесовские классификаторы, информационный пузырь, корректирующие коэффициенты, критический ресурс

RELEVANCE ADJUSTMENT METHOD FOR RECOMMENDATION SYSTEMS

¹Mitrokhin M.A., ²Martyshkin A.I., ¹Ionova D.N.,
¹Fedyashov M.S., ¹Gorynina A.V.

¹*Penza State University, Penza,*

e-mail: mmax83@mail.ru, ionova243@gmail.com, mark02042001@mail.ru, zmey.6453@mail.ru;

²*Penza State Technological University, Penza, e-mail: mai@penzgtu.ru*

The article discusses a method of relevance adjustment when compiling a list of recommendations, taking into account the need to synchronize access to a critical resource during asynchronous operation. The sensitivity of various recommendation algorithms based on user preferences to such problems as cold start, missing values, information bubble was analyzed. A group of data-based approach methods were collaborative filtering, content-based filtering, and knowledge-based filtering. Bayesian classifiers, matrix factorization, and neural networks were considered as examples of model-based approach methods. The method based on Bayesian classifiers has become preferable for further consideration due to the speed of implementation and simplified design process. The method of solving the problem of cold start and information bubble for the Bayesian recommendation algorithm based on corrective indicators is described: negative weights, attenuation coefficient and novelty coefficient, which make it possible to diversify the user's selections. In addition, a mechanism was proposed for synchronizing the process of changing the array of negative weights, which is a critical resource provided that corrective indicators are used together. As a visual representation of the mechanism for editing an array of negative weights, the method of formalization using event-driven nondeterministic automata was used.

Keywords: recommendation systems, Bayesian classifiers, information bubble, correction factors, critical resource

С увеличением объема данных в Интернете пользователю становится непросто своевременно найти интересующую информацию. Рекомендательные системы – это программные компоненты, предсказывающие на основании данных о пользователе или его предыдущих действиях объекты, которые заинтересуют пользователя в дан-

ный момент времени. Такими объектами могут стать фильмы, книги, товары, новости, услуги, веб-сайты.

Рекомендательные системы решают следующие задачи: сбор данных об объектах интереса, формирование оценки объекта интереса и выдача рекомендаций пользователю. В настоящее время используются

несколько подходов к составлению рекомендаций на основе машинного обучения [1, с. 40]. В каждом из применяемых методов существуют свои особенности, неминуемо приводящие к проблемам.

Материалы и методы исследования

К частым проблемам методов составления рекомендаций можно отнести:

- холодный старт (*cold start*);
- пропущенные значения или разреженность данных (*missing values*);
- информационный пузырь (*freshness in recommendations*).

Проблема холодного старта проявляется тогда, когда невозможно дать надежные прогнозы ввиду недостаточности или полного отсутствия входных данных о пользователе или объекте интереса [2, с. 28].

Частными случаями являются:

- регистрация нового пользователя в систему, о котором на данный момент ничего не известно и который еще ничего просмотрел [3, с. 54];
- добавление новых, еще никем не оцененных товаров или услуг.

Только что зарегистрированные пользователи не получают персонализированные рекомендации на основе предпочтений и могут уйти из информационной системы, не найдя в короткий срок релевантную информацию. Исходя из этого, сервисы должны предоставлять потребителям на момент регистрации первичную информацию о товарах и услугах, а после получения отклика предоставлять релевантные рекомендации.

Проблема холодного старта рекомендательной системы также возникает при недостаточном количестве откликов пользователей на объекты интереса. Для решения данной проблемы информационные системы должны либо поощрять потребителей за оценки товаров или услуг, либо использовать иные методы прогнозирования, базирующиеся на неявном взаимодействии пользователя с системой.

Проблема холодного старта свойственна коллаборативной фильтрации, поскольку данному методу для корректного функционирования необходима минимальная база знаний о взаимодействиях пользователей с объектами интереса.

Проблема пропущенных значений возникает из-за разреженности информации в базах данных рекомендательных систем. Это подразумевает, что пользователи при взаимодействии с сервисом не оценивают объекты интереса, поскольку данный товар им не интересен или они не знают о существовании альтернативных продуктов, подходящих под их потребности. Вы-

яснение причин возникновения неопределенных значений атрибутов помогает увеличить эффективность персонализированного прогнозирования, в противном случае велик риск снижения точности рекомендаций, а также возникновения неоднозначной информации [4, с. 132].

Проблема разреженности данных встречается в коллаборативной фильтрации и методе рекомендаций на основе контента, так как из-за большого количества неопределенных данных происходит смещение оценок в их сторону, что увеличивает погрешность прогнозирования. Одним из простейших способов урегулирования данной проблемы является замена неопределенных параметров ненулевыми показателями. К другому способу можно отнести кластеризацию данных, при которой осуществляется разделение задач на подзадачи с наиболее связанной информацией внутри каждого кластера.

Проблема информационного пузыря рекомендаций свойственна в основном крупным информационным системам с постоянно пополняющейся базой объектов интереса и с неизменяемыми предпочтениями субъектов. При этом пользователю, повторно обратившемуся к системе с аналогичным запросом, постоянно предоставляется одна и та же подборка наиболее релевантных объектов, что приводит к утрате интереса к целевой платформе. Данная проблема встречается в алгоритмах, основанных на моделях, поскольку процесс их переобучения является длительным и трудоемким. Исходя из этого, существует необходимость использования дополнительных алгоритмов и корректирующих коэффициентов, позволяющих осуществлять подбор рекомендаций вне полного цикла обучения модели. Решение данного вопроса позволит разнообразить релевантные подборки, добавляя в них не просмотренные ранее или совершенно новые объекты интереса.

Существенным недостатком алгоритмов фильтрации на основе знаний и нейронных сетей является сложность в проектировании, заключающаяся в больших затратах на реализацию, поскольку возникают необходимость в детальной проработке предметной области и потребность в сборе обширной базы знаний о предметной области, в которой применяется система [5, с. 3]. Матричная факторизация, как и нейронные сети, требует переобучения модели при поступлении новых исходных данных, что замедляет процесс прогнозирования, а это, в свою очередь, негативно влияет на скорость реакции рекомендательной системы в целом.

Таблица 1

Сравнение методов подхода на основе данных по их чувствительности к проблемам рекомендательных механизмов

Проблема / Метод	Коллаборативная фильтрация	Фильтрация на основе контента	Фильтрация на основе знаний
Холодный старт	+	+	–
Пропущенные значения	+	+	–
Проблема информационного пузыря	+	+	+

Таблица 2

Сравнение методов подхода на основе моделей по их чувствительности к проблемам рекомендательных механизмов

Проблема / Метод	Матричная факторизация	Нейронные сети	Байесовские классификаторы
Холодный старт	–	–	–
Пропущенные значения	–	–	–
Проблема информационного пузыря	+	+	+

На основе описанных выше проблем можно сделать вывод, что на данный момент не существует универсального алгоритма составления рекомендаций [6, с. 19]. Наличие (+) или отсутствие (–) конкретных проблем в рассматриваемых методах отражено в таблицах 1, 2.

Предпочтительными среди рассмотренных ранее алгоритмов являются байесовские классификаторы за счет скорости в реализации и упрощенного процесса проектирования. Байесовские классификаторы менее требовательны к объему исходных данных для обучения модели и обладают высоким уровнем масштабируемости, кроме того, они достаточно просты и эффективны при работе с большим массивом входных данных.

Предлагаемый подход к корректировке релевантности формирования прогнозов

Байесовские классификаторы представляют собой вероятностную модель, где атрибуты объектов и метка класса рассматриваются как дискретные величины. Он запоминает все признаки, а также вероятности того, что данный атрибут входит в конкретную классификацию (класс). Набор данных поступает в классификатор последовательно, после выполнения проверки наличия атрибутов в классе происходит обновление массива объектов интереса [7, с. 30; 8, с. 3]. Таким способом происходит обучение контентной модели.

Вероятностный классификатор предсказывает класс C_i с самой большой условной

вероятностью P_i для заданного вектора признаков $x = (x_1 \dots x_n)$:

$$P(C_i | x_1 \& \dots \& x_n) = \frac{P(x_1 | C_i) \dots P(x_n | C_i) P(C_i)}{P(x_1) \dots P(x_n)} \quad (1)$$

Вероятность класса C_i находится на основе параметров $(x_1 \dots x_n)$, подразумевается, что признаки являются независимыми, следовательно, приведенная выше вероятность пропорциональна:

$$P(C_i) \prod P(x_i | C_i). \quad (2)$$

Исходя из этого, байесовские классификаторы можно рассматривать как функцию, которая каждому выходному значению модели присваивает метку класса, т.е. $y = C_i$, следующим образом:

$$y = \operatorname{argmax} \prod P(x_i | C_i). \quad (3)$$

Из этого следует, что условная вероятность P_i , вычисленная для каждого класса C_i и признака x_i , присваивается к имеющему большую вероятность классу C_i . Результатом обучения контентной модели является список атрибутов вместе с условными вероятностями.

Несмотря на предположение, что признаки являются независимыми, байесовские классификаторы обеспечивают высокую точность классификации, решая проблемы холодного старта и пропущенных значений, однако нерешенным остается

ся вопрос информационного пузыря рекомендаций [9, с. 57]. Для этого необходимо использовать дополнительные показатели, позволяющие адаптировать список прогнозов под особенности конкретной рекомендательной системы.

Первоначально список рекомендаций составляется на основании общей статистики популярности объектов интереса, которая, в свою очередь, включает в себя предпочтения большинства пользователей системы. В дальнейшем формирование релевантной подборки происходит в рамках цикла обратной связи, представленного на рисунке 1, который включает в себя следующие этапы:

- 1) составление рекомендаций;
- 2) взаимодействие пользователя с системой;
- 3) составление рекомендаций на основе интересов пользователя.

На первом этапе система выдает подборку объектов интереса по соответствующему запросу, исходя из существующей информации о пользователе и объектах интереса.

Второй этап включает в себя действия текущего пользователя в целевой системе, а именно:

- поиск объекта интереса;
- переход на страницу объекта интереса;
- просмотр детализации характеристик объекта интереса;
- добавление объекта интереса в избранное;
- бронирование / добавление в корзину объекта интереса;
- покупка объекта интереса.

На третьем этапе данного цикла происходит обновление списка рекомендаций на основе предпочтений пользователя, которые были выявлены на предыдущем этапе.

В рамках цикла обратной связи при обновлении рекомендаций необходимо учитывать не только непосредственные действия пользователя, но и дополнительные коэф-

фициенты, формирующиеся на основании неявных процессов, происходящих в системе, поскольку они позволяют повысить уровень точности прогнозов. К подобным показателям относятся:

- коэффициент новизны;
- отрицательные веса;
- коэффициент затухания.

Коэффициент новизны применяется на заключительной стадии составления подборки рекомендаций для новых объектов интереса, имеющих низкую оценку релевантности ввиду отсутствия информации о них (формула 4):

$$q(k) = y + l_k, \quad (4),$$

где k – текущая информация об объекте интереса относительно пользователя;

y – список атрибутов вместе с условными вероятностями, вычисление которых происходит по формуле 3;

l_k – текущее значение коэффициента новизны данного объекта интереса.

Этот механизм позволит добавить продукт в список прогнозов наряду с наиболее подходящими под интересы пользователя продуктами. При инициализации нового объекта в целевой системе коэффициенту новизны присваивается максимальное значение, после чего в течение n дней он декрентируется до минимума.

Под отрицательными весами понимается массив данных, содержащий в себе информацию о степени заинтересованности пользователя в данном объекте интереса. Коэффициент, значение которого обратно пропорционально уровню востребованности текущего объекта интереса у пользователя, применяется на заключительной стадии расчета оценки релевантности (формула 5):

$$q(k) = y + d_k, \quad (5),$$

где d_k – текущее значение отрицательного веса данного объекта интереса.

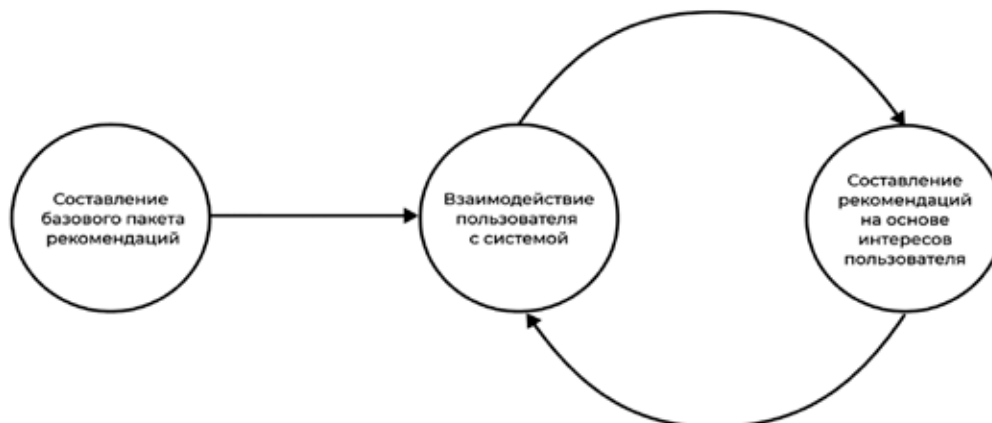


Рис. 1. Формирование релевантной подборки в рамках цикла обратной связи

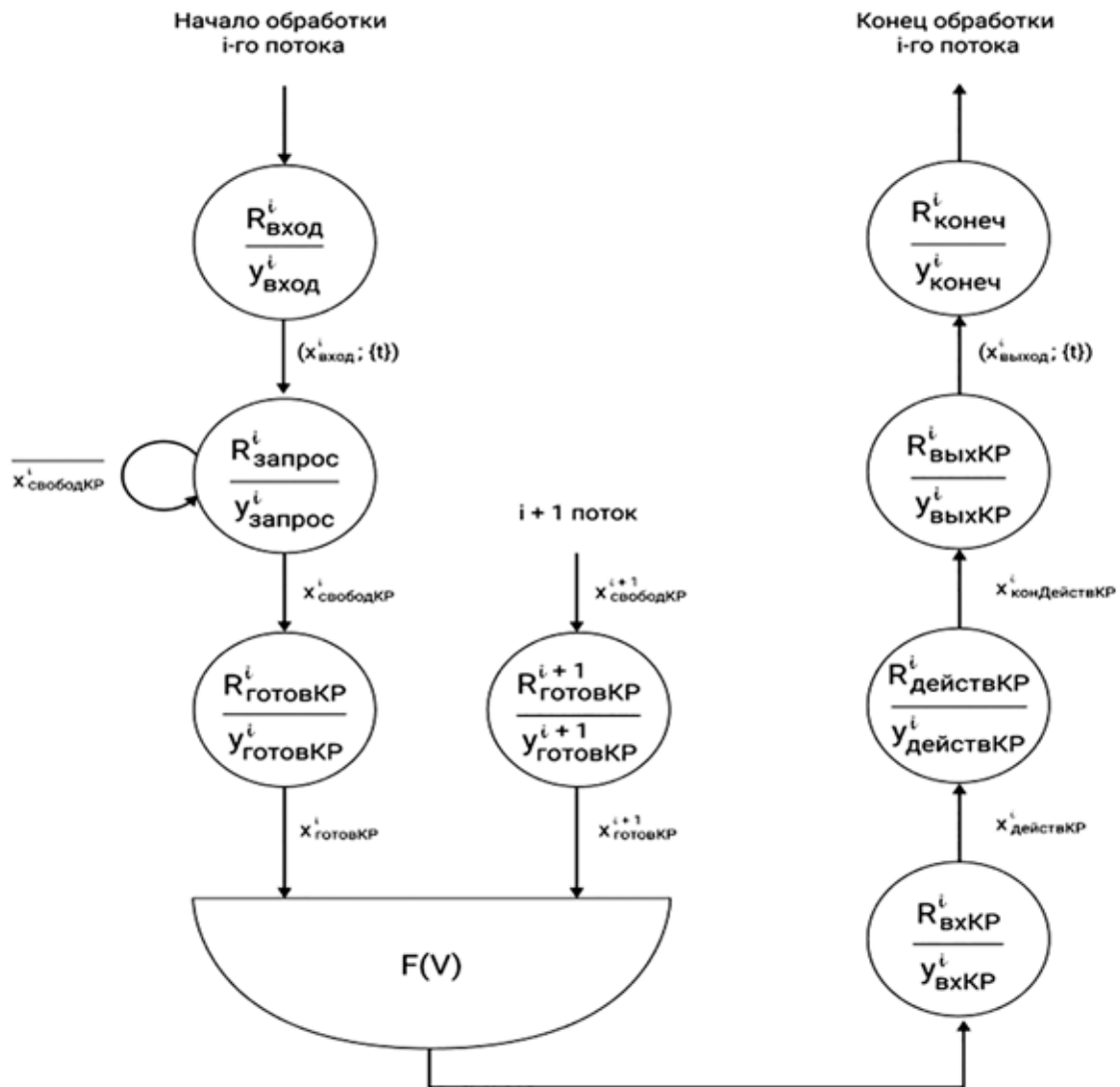


Рис. 2. Математическая модель изменения массива отрицательных весов

Формирование массива отрицательных весов происходит на втором этапе цикла обратной связи и зависит от явных поведений пользователя, представленных ранее, и неявных процессов системы, например от времени ожидания следующего действия субъекта относительно рекомендуемого объекта после перехода на его страницу. В данном случае считается, что коэффициент отрицательного веса для текущего продукта увеличивается, если субъект по истечении некоторого времени не добавляет объект в избранное/корзину или не бронирует его, поскольку это может интерпретироваться как отсутствие интереса к данному продукту после изучения его характеристик. Однако возможно возникновение проблемы неверного подсчета коэффициента ввиду работы с объектом несколькими пользователями.

Таким образом, продукты, имеющие высокий показатель отрицательного веса, оказываются в конце списка прогнозов или же вообще не попадают в него. Однако данный продукт, несмотря на высокий показатель отрицательного веса, может попасть в список рекомендаций в случае превышения оценки релевантности над отрицательным коэффициентом и даже находиться на высоких позициях до тех пор, пока действия пользователя не снизят его востребованность до критического уровня.

Подобные рекомендации, имеющие высокий отрицательный вес и низкую оценку релевантности, всегда будут находиться в конце списка прогнозов. Исходя из этого, пользователю, повторно обратившемуся к системе с аналогичным запросом, будет предоставляться одна и та же подборка

продуктов, наиболее соответствующих интересам субъекта. Во избежание подобных ситуаций применяется корректирующий механизм – коэффициент затухания, позволяющий вернуть объект интереса, имеющий высокий отрицательный вес, в подборку рекомендаций через определенный период времени. Коэффициент затухания может быть использован несколькими способами:

- изменение отрицательного веса для данного объекта до минимального значения спустя n дней после инициализации исходного значения;

- изменение отрицательного веса для данного объекта до минимального значения в случае, если этот продукт продолжает быть релевантен, но не попадает в подборку из-за высокого показателя отрицательного веса более чем k раз;
- изменение массива отрицательных весов для данной подборки до минимального значения спустя выполнения k рекомендаций;
- изменение массива отрицательных весов для данной подборки до минимального значения спустя n дней после инициализации исходного массива.

Таблица 3

Условные обозначения состояний автомата

Состояние	Входной сигнал	Выходной сигнал
$R^i_{\text{вход}}$ – состояние, инициализирующее начало работы i -потока с массивом отрицательных весов	–	$y^i_{\text{вход}}$ – i -й поток инициализирован
$R^i_{\text{запрос}}$ – состояние, определяющее доступность критического ресурса для i -го потока	$x^i_{\text{вход}}$ – сигнал отправки запроса на доступ к критическому ресурсу $!x^i_{\text{свободКР}}$ – сигнал ожидания освобождения критического ресурса	$y^i_{\text{запрос}}$ – доступность критического ресурса
$R^i_{\text{готовКР}}$ – состояние, определяющее готовность к входу i -го потока в критический ресурс	$x^i_{\text{свободКР}}$ – сигнал доступности критического ресурса i -ому потоку	$y^i_{\text{готовКР}}$ – готовность критического ресурса к обработке i -м потоком
$R^{i+1}_{\text{готовКР}}$ – состояние, определяющее готовность i -го+1 потока на обслуживание	$x^{i+1}_{\text{свободКР}}$ – сигнал критического ресурса к обработке $i+1$ потоком	$y^{i+1}_{\text{готовКР}}$ – готовность критического ресурса к обработке $i+1$ потоком
$R^i_{\text{вхКР}}$ – состояние, определяющее вход i -го потока в критический ресурс	$x^i_{\text{готовКР}} \vee x^{i+1}_{\text{готовКР}}$ – сигнал входа потока в критическую секцию	$y^i_{\text{вхКР}}$ – i -й поток вошел в критическую секцию
$R^i_{\text{действияКР}}$ – состояние, определяющее действия над i -м потоком внутри критической секции	$x^i_{\text{действияКР}}$ – сигнал начала обработки критического ресурса i -м потоком	$y^i_{\text{действияКР}}$ – критический ресурс обработан i -м потоком
$R^i_{\text{выхКР}}$ – состояние, свидетельствующее о выходе i -го потока из критического ресурса (освобождение критической секции)	$x^i_{\text{выхКР}}$ – сигнал выхода i -го потока из критической секции	$y^i_{\text{выхКР}}$ – критический ресурс освобожден
$R^i_{\text{конец}}$ – конечное состояние, свидетельствующее о завершении работы с массивом отрицательных весов	$x^i_{\text{выход}}$ – сигнал завершения работы с массивом отрицательных весов	$y^i_{\text{конец}}$ – работа с массивом отрицательных весов завершена

Исходя из этого, применение коэффициента затухания вместе с отрицательными весами позволит частично решить проблему информационного пузыря рекомендаций, добавив в подборки не просмотренные ранее или заведомо не актуальные объекты интереса.

Необходимо принять во внимание, что совместное применение отрицательных весов и базирующегося на них коэффициента затухания может привести к неверному составлению прогноза для конкретного продукта, поскольку существует вероятность одновременного изменения значения отрицательного показателя во время явных действий пользователя с текущим объектом и непосредственным применением коэффициента затухания. Исходя из этого, необходимо разработать механизм доступа к критическому ресурсу, которым будет являться массив отрицательных весов.

Алгоритм доступа к массиву можно формализовать при помощи конечного автомата, что отражено на рисунке 2 [10, с. 8]. Применение данного математического аппарата позволит наглядно представить механизмы взаимодействия с критическим ресурсом и решения тупиковых ситуаций, возникающих при совместной работе нескольких потоков с массивом отрицательных весов.

Условные обозначения данного механизма представлены в таблице 3.

Начальным состоянием, инициализирующим работу i -го потока с критическим ресурсом, являются явные действия пользователя с текущим объектом интереса или непосредственное применение коэффициента затухания [10]. На данном этапе критический ресурс еще не готов к обслуживанию текущим потоком. Для этого необходимо запросить доступность объекта синхронизации; если он свободен, происходит переход к состоянию готовности, в противном случае поток ожидает освобождение ресурса. В случае захвата потом критического ресурса ему предоставляется возможность редактирования массива отрицательных весов, после завершения которого поток освобождает объект синхронизации и переходит в конечное состояние.

Применение байесовских классификаторов совместно с корректирующими коэффициентами позволяет создать универсальный механизм прогнозирования, решающий проблемы рекомендательных систем, и адаптирует список релевантных объектов интереса под особенности конкретной целевой платформы.

Заключение

В данной статье были рассмотрены задачи, которые решают рекомендательные системы, проанализированы проблемы составления рекомендаций на основе предпочтений пользователей в существующих методах прогнозирования и существующие подходы к их решению, наиболее актуальной из них стала проблема информационного пузыря рекомендаций. Для решения данного вопроса был предложен механизм, сочетающий в себе байесовские классификаторы и корректирующие показатели: отрицательные веса, коэффициент затухания и коэффициент новизны. Кроме того, был предложен механизм синхронизации процесса изменения массива отрицательных весов, необходимый для исключения вероятности одновременного изменения значения отрицательного показателя во время явных действий пользователя с текущим объектом и непосредственным применением коэффициента затухания. Механизм представлен с использованием метода формализации с применением событийных недетерминированных автоматов. Данный подход к решению проблем рекомендательных систем позволит разнообразить подборки объектов интереса, добавляя в них совершенно новые или не просмотренные ранее.

Список литературы

1. Гудфеллоу Я., Бенджио И. Глубокое обучение. М.: ДМК Пресс, 2018. 652 с.
2. Chen A. The Cold Start Problem. HarperCollins Publishers Ltd, 2021. 198 p.
3. Gilman E. Zero Trust Networks: Building Secure Systems in Untrusted. O'Reilly, 2017. 125 p.
4. Grus J. Data Science from Scratch, 2nd Edition. O'Reilly, 2019. 398 p.
5. Charu C. Aggarwal Data Mining: The Textbook. Springer International Publishing, 2015. 734 p.
6. Королева Д.Е., Филиппов М.В. Анализ алгоритмов обучения коллаборативных рекомендательных систем // Инженерный журнал: наука и инновации. 2013. №6 (18). С. 23.
7. Звягин Л.С. Метод байесовских сетей и ключевые аспекты байесовского моделирования // Международная конференция по мягким вычислениям и измерениям. СПб.: Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина), 2018. С. 30-34.
8. Cheng J., Greiner R. Comparing Bayesian Network Classifiers. 2013. P. 101-108. DOI: 10.48550/arXiv.1301.6684.
9. Литвиненко Н.Г., Литвиненко А.Г., Мамырбаев О.Ж., Шахметова А.С. Байесовские сети. Теория и практика. Алматы: Институт информационных и вычислительных технологий КН МОН РК, 2020. 197 с.
10. Волчихин В.И., Вашкевич Н.П., Бикташев Р.А. Модели событийных недетерминированных автоматов представления алгоритмов управления взаимодействующими процессами в многопроцессорных вычислительных системах на основе использования механизма монитора // Известия вузов. Поволжский регион. Технические науки. 2013. № 2(26). С. 5-14.