

СПОСОБЫ ПОСТРОЕНИЯ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ

METHODS FOR BUILDING
RECOMMENDER SYSTEMSYu. Morozov
S. Saradgishvili

Summary. Recommender systems are programs that predict which objects are of interest to a given system. Recommender systems have increasingly become part of modern Internet services and online stores. This paper provides an overview of approaches to the formation of attention in the recommendatory principles.

Keywords: recommender systems; methods of recommender systems; collaborative filtering; recommendation methods; content methods; data analysis.

Морозов Юрий АлексеевичАспирант, Санкт-Петербургский политехнический
университет Петра Великого
stonefiz@gmail.com**Сараджишвили Сергей Эрикович**К.т.н., доцент, Санкт-Петербургский
политехнический университет Петра Великого
ssaradg@yandex.ru

Аннотация. Рекомендательные системы — это программы, которые пытаются предсказать какие объекты интересны пользователю данной системы. Рекомендательные системы всё больше становятся частью современных интернет-сервисов и интернет-магазинов. В данной работе проводится обзор подходов к формированию рекомендаций в рекомендательных системах.

Ключевые слова: рекомендательные системы; методы рекомендательных систем; коллаборативная фильтрация; методы рекомендаций; контентные методы; анализ данных.

Рекомендательные системы — это информационные системы для прогноза различных вариантов объектов, товаров или услуг конкретному пользователю [1]. Рекомендации пользователю каких-либо объектов основываются на данных, которые могут быть получены различными способами и будут связаны косвенно или напрямую с пользователем и этими объектами. В качестве данных для анализа могут быть получены через оценку пользователем похожих объектов предварительно (например, оценка понравившихся товаров в интернет-магазинах), могут быть взяты количества прослушиваний музыкальных композиций в музыкальных интернет-сервисах, а в дальнейшем ранжированы по количеству их прослушиваний. Также могут быть взяты количественные характеристики просмотренных объектов в видео-сервисах — самые часто просматриваемые пользователем фильмы и сериалы и т.д.

Несмотря на то, что рекомендательные системы пытаются достичь одной и той же цели, для неё они используют разные методы. В методах построения рекомендательных систем выделяют следующие категории методов:

- ◆ методы коллаборативной (совместной) фильтрации;
- ◆ методы, основанные на контенте (контентные методы);

- ◆ методы, основанные на знании потребностей клиентов (методы, основанные на знаниях);
- ◆ гибридные методы

Коллаборативная фильтрация — это популярный метод предсказания объектов или информации для пользователя системы, который работает с использованием базы данных, которая в свою очередь отражает предпочтения пользователями различных элементов, также называется матрица пользовательских элементов. [2] На рисунке ниже представлен пример матрицы пользовательских элементов, на которых строится коллаборативная фильтрация.

Для данного подхода справедливо предположение, что люди, которые были согласны с чем-то в прошлом, будут согласны с этим и в будущем, и в результате им будут нравиться подобные типы элементов, которые им также нравились в прошлом. Примером данного типа фильтрации может быть такой случай, как мне и моему другу в прошлом нравился одинаковый набор книг, а ему нравится книга, которую я не читал, но потому что мы в прошлом нам нравилась одна и та же книга, то опираясь на данный подход новая книга, которую я не читал, мне тоже должна будет понравиться, таким образом рекомендательная система предложит мне её. Основываясь на этом примере, вместо того чтобы отталкивать

	Item ₁	Item ₂	Item _j	Item _n
User ₁					
User ₂					
.					
User _i					
User _m					

Рис. 1. Матрица пользовательских элементов

ся только от того что нравится моему другу, мы можем также отталкиваться от множества товаров и объектов, которые понравились ранее разным пользователям, и рекомендовать новый товар на основе сходства между этими объектами, которые понравились разным людям. Тут и рассчитывается сходство с использованием оценок и матриц пользовательских элементов. Логика этого алгоритма известна как совместная фильтрация на основе элементов.

Контентные методы фильтрации рекомендательных систем основаны на анализе характеристик и описания товаров и покупателей. Ключевой идеей контентной фильтрации является предположение, что покупатель заинтересован в товарах, аналогичных тем, что куплены. Когда нужно рекомендовать такие документы, как веб-страницы, новости и публикации, метод фильтрации на основе содержимого является наиболее успешным. Т.е. это лучше всего подходит в ситуациях, когда известна информация об элементе, но мало известно информации о пользователе.

Данный подход, основан на описании и характеристиках объектов, которые связаны с пользователем или он покупает эти объекты. Такой подход предполагает наличие большого количества данных о товарах и покупателях (или об объектах и пользователях). Например, все покупатели при приобретении товара заполняют анкету, где они описывают свои социальные и личные характеристики, интересы и т.д. Например, бренд, назначение, цвет и цена товаров могут быть использованы в качестве их описания в интернет-магазинах. На основе накопленных данных о взаимодействиях между покупа-

телями и товарами можно построить выборку для обучения моделей нейронных сетей, чтобы в дальнейшем прогнозировать рекомендуемые товары для пользователя. Также помимо нейронных сетей можно использовать другие различные подходы — модели векторного пространства, такие как частота терминов, обратная частоте документа (TF/IDF), использовать наивно-байесовский классификатор или другие методы статистического анализа. Одним из важных моментов в методах контентной фильтрации можно отметить, что они не нуждаются в профилях других пользователей, поскольку они не влияют на рекомендацию, в отличие от методов коллаборативной фильтрации.

Методы, основанные на знаниях, которые используют знания о потребностях покупателя, служат для построения рекомендуемых экспертами систем. В таких системах алгоритм учитывает знания об объектах и элементах, о пользовательские предпочтения, указанных явно, и разработанных критериях рекомендаций, прежде прогнозировать рекомендуемые объекты. [3] Рекомендации в таких системах, как правило формируются на основе сформулированных вручную правил ассоциаций. Например, мы можем взять ситуацию, при которой создается система, рекомендуемая бытовую технику, такую как кондиционеры для пользователей. В этом случае система должна будет учитывать особенности и характеристики предметов, а пользовательские данные будут собираться путем получения от пользователей дополнительной информации и пожеланий по требованиям характеристик, после чего на основе собранных данных на основе правил по полученным предпочтениям пользователей будут выдаваться рекомендации.

Природа предметной области может определять как синтаксис, так и подходы к разработке таких информационных систем. Для областей, которые имеют серьезные последствия при плохих рекомендациях, система должна быть основана на сильных исследованиях и может быть императивной. Для областей, в которых не может быть серьезных последствий, может быть достаточно системы, основанной на неподтвержденном профессиональном опыте.

В целом, такой подход позволяет обеспечить очень высокое качество рекомендаций, хотя разработка таких систем требует привлечения квалифицированных специалистов для составления, обновления, проверки и добавления правил ассоциаций.

Гибридные методы для построения рекомендательных систем гораздо более распространены в реальном мире, поскольку объединение компонентов из различных подходов может преодолеть различные недостат-

ки. Идея такого подхода заключается в том, что комбинация различных подходов и алгоритмов будет давать более точные и эффективные рекомендации, чем один алгоритм, поскольку слабые стороны одного алгоритма могут быть компенсированы другим алгоритмом. Комбинация подходов может осуществляться любым из способов: отдельная реализация алгоритмов или объединение результатов, использование контентной фильтрации в коллаборативной фильтрации, использование коллаборативной фильтрации в контентном подходе, в общем системы рекомендаций, которая использует вместе разные подходы.

Рекомендательные системы можно с уверенностью назвать перспективным направлением, открывающим новые возможности для получения новых данных пользователями сети, а также несущие особенную ценность для интернет-сервисов, предоставляющих различные услуги, интернет-магазинов и интернет-бизнеса в целом.

ЛИТЕРАТУРА

1. Цурко В.В. Рекомендательные системы в здравоохранении, «Управление в медико-биологических и экологических системах», Москва — 2019
2. Jurij A. Morozov; Sergej E. Saradgishvili Improving Collaborative Filtering — 2021 IV International Conference on Control in Technical Systems (CTS) — 2021
3. Aditeya Pandey, Sehi L'Y, Nils Gehlenborg Towards a Knowledge-Based Recommendation System for Genomics Visualization — OSF — 2020
4. Francesco RicciLior RokachBracha ShapiraPaul B. Kantor Recommender Systems Handbook — Springer — 2011

© Морозов Юрий Алексеевич (stonefiz@gmail.com), Сараджишвили Сергей Эрикович (ssaradg@yandex.ru).
Журнал «Современная наука: актуальные проблемы теории и практики»