

УДК 004.89

*Рытиков А.В., студент 1 курса магистратуры, гр. ИСТмз-21
ФГОБУ ВО «Поволжский государственный университет
телекоммуникаций и информатики», Россия, г. Самара*

ОПИСАНИЕ РАБОТЫ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ НА ОСНОВЕ СОВМЕСТНОЙ ФИЛЬТРАЦИИ И КОНТЕНТ ОРИЕНТИРОВАННЫХ МЕТОДАХ

Аннотация. В статье рассматриваются основные методы работы рекомендательных систем, такие как метод совместной фильтрации и контент ориентированные методы. Для каждой модели описаны общие принципы работы, их теоретические основы, а так же сильные и слабые стороны методов.

Ключевые слова: рекомендательные системы, методы совместной фильтрации, контент ориентированные методы.

*Rytikov A.V., 1st year Master's student,
of gr. ISTMZ-21 FGOBU in « Povolzhskiy State University of
Telecommunications and Informatics », Russia, Samar*

DESCRIPTION OF THE WORK OF RECOMMENDATION SYSTEMS BASED ON JOINT FILTERING AND CONTENT- ORIENTED METHODS

Abstract. The article discusses various paradigms of recommendation systems. For each of the paradigms, we will present the general principles of operation, describe their theoretical foundations and discuss their strengths and weaknesses.

Keywords: recommendation systems, collaborative filtering methods, content-oriented methods.

В современное время, системы рекомендаций стали занимать все больше места в нашей жизни. Начиная с электронной коммерции и

заканчивая рекламой в Интернете, рекомендательные системы сегодня неизбежны в нашей ежедневной онлайн-жизни [1].

Рекомендательные системы - это алгоритмы, предназначенные для предложения пользователям соответствующих предметов (текста для чтения предметов, которые можно смотреть в фильмах, продуктов для покупки и многое другое, в зависимости от отрасли) [2].

Рекомендательные системы очень важны в некоторых отраслях, так как они могут приносить огромную прибыль, когда они эффективны, или также могут значительно отличаться от конкурентов.

Основная цель системы рекомендаций - предлагать пользователям соответствующие элементы. Существуют две основные категории методов для решения подобной задачи: методы совместной фильтрации и методы, основанные на контенте.

Совместные методы фильтрации для рекомендательных систем - основаны исключительно на прошлых взаимодействиях, зарегистрированных между пользователями и предметами, с целью выработки новых рекомендаций. Основная идея, заключается в том, что прошлые взаимодействия пользователь-элемент достаточно для того, чтобы обнаружить аналогичных пользователей и / или аналогичных элементов и прогнозирования на основе этих предполагаемых приближений.

Класс алгоритмов совместной фильтрации разделен на две подкатегории, которые обычно называются основанными на памяти и модельными подходами. Подходы, основанные на памяти, напрямую работают со значениями записанных взаимодействий, предполагая отсутствие модели, и по существу основаны на поиске ближайших соседей (например, находят ближайших пользователей от интересующего пользователя и предлагают наиболее популярные элементы среди этих соседей). Подходы, основанные на моделях, предполагают базовую

«порождающую» модель, которая объясняет взаимодействие пользователя с элементом и пытается обнаружить ее, чтобы делать новые предсказания.

Основное преимущество совместных подходов заключается в том, что они не требуют информации о пользователях или элементах и, следовательно, могут использоваться во многих ситуациях. Более того, чем больше пользователей взаимодействуют с элементами, тем больше новых рекомендаций становятся точными: для фиксированного набора пользователей и элементов новые взаимодействия, записанные с течением времени, приносят новую информацию и делают систему более эффективной.

Подходы на основе контента используют дополнительную информацию о пользователях и / или элементах. Для примера возьмем системы рекомендации фильмов, такой дополнительной информацией может быть, например, пол, возраст, работа или любая другая личная информация для пользователей, а также категория, главные действующие лица, продолжительность или другие характеристики.

Идея методов, основанных на контенте, состоит в том, чтобы попытаться построить модель, основанную на доступных «функциях», которые объясняют наблюдаемые взаимодействия пользователя с элементом. Методы, основанные на контенте, гораздо меньше страдают от проблемы «холодного запуска», чем совместные подходы: новые пользователи или элементы могут быть описаны по их характеристикам (контенту), и поэтому могут быть сделаны соответствующие предложения для этих новых объектов. Этому недостатку логически будут страдать только новые пользователи или элементы с ранее невиданными функциями, но как только система станет достаточно старой, у нее будет мало шансов вообще не случиться.

В основанных на памяти совместных методах скрытая модель не предполагается. Алгоритмы напрямую работают с взаимодействиями

пользователь-элемент: например, пользователи представлены своими взаимодействиями с элементами, и поиск ближайших соседей по этим представлениям используется для создания предложений. Поскольку скрытая модель не предполагается, эти методы теоретически имеют низкое смещение, но высокую дисперсию.

В основанных на модели методах сотрудничества предполагается некоторая модель скрытого взаимодействия. Модель обучена восстанавливать значения взаимодействий пользователь-элемент из собственного представления пользователей и элементов. Новые предложения могут быть сделаны на основе этой модели. Скрытые представления пользователей и элементов, извлеченные моделью, имеют математическое значение, которое может быть трудно интерпретировать для человека. Поскольку предполагается (довольно бесплатная) модель взаимодействия пользователя с элементом, эти методы теоретически имеют более высокий уклон, но более низкую дисперсию, чем методы, предполагающие отсутствие скрытой модели [3].

Наконец, в методах, основанных на контенте, также предполагается некоторая модель скрытого взаимодействия. Однако здесь модель снабжена контентом, который определяет представление пользователей и / или элементов: например, пользователи представлены заданными функциями, и мы пытаемся смоделировать для каждого элемента тип профиля пользователя, который нравится этому элементу или нет. Здесь, что касается основанных на модели методов взаимодействия, предполагается модель взаимодействия пользователя с элементом. Однако эта модель более ограничена (поскольку дано представление пользователей и / или элементов), и поэтому метод имеет тенденцию иметь самое высокое смещение, но самую низкую дисперсию.

Использованные источники:

1. Пройдл А. История Систем рекомендаций // Телеспутник, 2015. № 3(233). С. 48-50.
2. Меньшикова Н. В., Портнов И. В., Николаев И. Е. Обзор рекомендательных систем и возможностей учета контекста при формировании индивидуальных рекомендаций // Academy. 2016. № 6. С. 20-22.
3. Игнатьев В. Ю., Лемтюжникова Д. В., Руль Д. И., Рябов И. Л. Построение гибридной рекомендательной системы с улучшением точности // Известия РАН. Теория и системы управления. 2018. № 6. С. 101-108.