Рекомендательная система представляет из себя комплекс программ и сервисов, которые анализируют предпочтения пользователей и пытаются их предсказать. Рекомендательные системы являются одним из самых распространённых примеров применения инструментов анализа данных. Они созданы для того, чтобы решить проблему избытка информации, с который сталкиваются пользователи.

Актуальность разработки рекомендательной системы связана, прежде всего, с тем, что она имеет ряд преимуществ как для пользователя, так и для бизнеса. Для пользователя это возможность получить продукт, который потенциально может ему понравится. Для бизнеса это возможность увеличить общее количество продаж, увеличить продажи разнообразных товаров, улучшить у пользователя общее впечатление о магазине.

Существует 3 типа рекомендательных систем, они основаны на различном понимании того, по каким признакам объект должен считаться предпочтительным.

1. Коллаборативная фильтрация.
2. Фильтрация на основе содержимого.
3. Системы, основанные на знаниях.

Коллаборативная фильтрация – это алгоритм прогнозирования, в котором, рекомендации стояться на основе предпочтений других пользователей, со схожими интересами.

Фильтрация на основе содержимого – это подход, в рамках которого близость объектов друг к другу определяется на основе содержимого этих объектов. Рекомендации в таких системах стояться от того, какие объекты понравились пользователю и какие объекты к ним близки.

В отличии от предыдущих подходов, рекомендации, основанные на знаниях, не зависят от оценки объекта или его описания, а основаны на более глубоких правилах для выявления объектов интереса.

Каждый из представленных выше подходов имеет свои преимущества и недостатки, для их выделения необходимо провести сравнительный анализ на основе ряда критериев:

1. Адаптивность.
2. Интуитивность.
3. Быстрота развёртывания.
4. Холодный старт.
5. Прозрачность.

Критерий адаптивность определяет, насколько система способна адаптироваться к изменениям. Системы использующие методы коллаборативной фильтрации или фильтрации на основе содержимого способны адаптироваться к новым данным. Однако же, системы, основанные на знаниях не способны быстро адаптироваться к новой информации.

Интуитивность представляет из себя некую случайность, то есть случайное нахождение релевантных объектов в том смысле, что эти объекты могут заинтересовать пользователя, даже если пользователь не инициировал такой поиск. Это свойство не может быть достигнуто в системах, основанных на содержании объектов, так как там не учитываются предпочтения пользователей. Однако же системы, основанные на знаниях способны воспроизвести эффект интуитивности, но его достижение очень сильно связано с опытом и уровнем инженера, который строит систему.

Холодный старт представляет из себя ситуацию, когда необходимо предоставить рекомендации, хотя данных об интересах пользователя никаких нет. Такая проблема встречается только в системах, использующих подход коллаборативной фильтрации.

Свойство прозрачности определяет, насколько система может обосновать, результат работы рекомендательного алгоритма. Системы, основанные на содержании, или алгоритмы коллаборативной фильтрации, опираются исключительно на данные, и объясняют все свои выводы только данными. В отличии от предыдущих подходов, методы основанные на званиях способны предоставить полное развёрнутое пояснение, почему получен именно такой результат.

|  |  |  |  |
| --- | --- | --- | --- |
| Подход | Коллаборативная  фильтрация | Фильтрация на  основе контента | Фильтрация на  основе знаний |
| Быстрое развертывание | Да | Да | Нет |
| Адаптивность | Да | Да | Нет |
| Интуитивность | Да | Нет | Нет |
| Холодный старт | Да | Нет | Нет |
| Прозрачность | Нет | Нет | Да |

Исходя из данных, полученных во время анализа сравнительных характеристик можно сделать ряд выводов. Подхода коллаборативной фильтрации, и методы, основанные на знаниях, следует использовать для создания рекомендаций с низкой степенью участия, например фильмы или музыка, или книги. Под низкой степенью участия понимается то, что составленная неправильная рекомендация не принесёт никаких серьёзных потерь. Системы же, основанные на знаниях, следует использовать, например в финансовой сфере или в задаче рекомендации квартир или машин. В таких системах, как правило гораздо меньше данных, что делает затруднительных использования первых двух подходов