

МАТЕМАТИКА РЕКОМЕНДАТЕЛЬНЫХ АЛГОРИТМОВ

Гомонов М.А., Васильев А.М. студенты

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Князюк Н.В. – канд. физ.-мат. наук, доцент

В работе рассмотрены четыре вида рекомендательных систем: контентная, коллаборативная, экспертная и комбинированная, изучена их математическая составляющая. Исследованы реальные примеры использования рекомендательных систем.

Рекомендательные системы – это технологии, используемые для предсказания и рекомендации продуктов или контента, наиболее подходящих для конкретного пользователя [1]. Принцип работы таких систем реализуется с помощью обучения искусственного интеллекта на векторах, содержащих характеристики объекта, по которым будут осуществляться поиск и фильтрация.

Контентная система строит рекомендации на основе уже просмотренного пользователем контента, подбирая объекты, наиболее похожие на оценённые [2, 3]. Пусть $I = \{i_1, \dots, i_m\}$ – множество объектов, $U = \{u_1, \dots, u_n\}$ – упорядоченное множество пользователей, $x_i \in R^d$ – вектор признаков объекта. Профиль пользователя (взвешенное среднее по оценённым объектам):

$$p_u = \frac{\sum_{i \in I_u} r(u, i) \cdot x_i}{\sum_{i \in I_u} r(u, i)},$$

где $r(u, i)$ – оценка пользователя u объекту i ; I_u – множество оценённых объектов. Рекомендации даются по степени сходства векторов, которую можно задать стандартным косинусным способом:

$$\cos(\theta) = \frac{p_u \cdot x_i}{|p_u| \cdot |x_i|}$$

Чем ближе $\cos(\theta)$ к 1, тем выше соответствие объекта вкусам пользователя. Среди плюсов данной системы можно выделить независимость от других пользователей, а среди недостатков – проблему “холодного старта”, то есть невозможность построить профиль пользователя без начальных оценок.

Система коллаборативной фильтрации предсказывает рекомендации на основе сходства предпочтений пользователей [4, 5]. Строится матрица оценок $R \in \mathbb{R}^{n \times m}$, где $r(u, i)$ – оценка пользователя u объекту i . Степень сходства двух пользователей определяется формулой:

$$\text{sim}(u, u') = \frac{\sum_{j=1}^m r(u, j) \cdot r(u', j)}{\sqrt{\sum_{j=1}^m (r(u, j))^2 \cdot \sum_{j=1}^m (r(u', j))^2}}, \quad (1)$$

где $r(u, j)$ – оценка пользователя u объекту j .

Все u' , у которых $\text{sim}(u, u') \geq \alpha$, где α – некоторое произвольное число из промежутка $[0, 1]$, вносим в список U . Тогда прогноз оценки для i -го объекта будет вычисляться по формуле:

$$p(u, i) = \frac{\sum_{u' \in U} r(u', i) \cdot \text{sim}(u, u')}{\sum_{u' \in U} \text{sim}(u, u')}. \quad (2)$$

На основе вычислений остаётся предсказать как можно больше оценок и рекомендовать пользователю наиболее релевантные из них. Достоинствами коллаборативной системы являются возможность выделять неочевидные совпадения вкусов пользователей и независимость от признаков объектов, среди минусов можно заметить проблему “холодного старта” как для пользователей, так и для объектов, а также смещение в сторону популярного контента.

Экспертная система построена на рекомендациях экспертов [6]. Она сравнивает нынешнего пользователя на соответствие признака i , в случае положительного результата, выдаёт ему вариант рекомендаций, полученный от эксперта ранее. Система сравнивает предпочтения пользователя с мнениями экспертов $E = \{e_1, \dots, e_k\}$, где e_i – вектор оценок эксперта. Тогда сходство пользователя с экспертом вычисляется по формуле (1) с учетом замены u' на e . Прогноз оценки рассчитывается аналогично формуле (2), однако суммирование ведётся по похожим экспертам. Среди преимуществ

системы можно выделить стандартизированные оценки экспертов, а среди недостатков – меньшую персонализацию.

Комбинированная система совмещает в себе плюсы всех рассмотренных систем для максимально точного подбора рекомендаций, но в то же время она сложнее в реализации и вычислительно дороже. Так как в каждой системе есть понятие вероятностей соответствия пользователю, в данной системе могут единым образом выбираться наиболее вероятные варианты [6].

На практике такие сервисы, как Яндекс Еда, Деливери Маркет, Самокат и Сбер Маркет, совмещают различные системы, но наиболее частым выбором становится именно комбинированная система [7].

Список использованных источников:

1. Рекомендательная система // Википедия: свободная энциклопедия [Электронный ресурс]. – Режим доступа: https://ru.wikipedia.org/wiki/Рекомендательная_система. – Дата доступа: 30.03.2026.
2. Контентная фильтрация и управление доступом к ресурсам сети интернет в образовательных учреждениях // КиберЛенинка [Электронный ресурс]. – Режим доступа: <https://cyberleninka.ru/article/n/kontentnaya-filtratsiya-i-upravlenie-dostupom-k-resursam-seti-internet-v-obrazovatelnyh-uchrezhdeniyah/viewer>. – Дата доступа: 30.03.2026.
3. Построение рекомендательных систем, часть 1 / ЛАНИТ // Хабр [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/companies/lanit/articles/420499>. – Дата доступа: 29.03.2026.
4. Коллаборативная фильтрация как метод учета интересов пользователей в рекламе продукции // КиберЛенинка [Электронный ресурс]. – Режим доступа: <https://cyberleninka.ru/article/n/kollaborativnaya-filtratsiya-kak-metod-uchyota-interesov-polzovateley-v-reklame-produktsii/viewer>. – Дата доступа: 29.03.2026.
5. Построение рекомендательных систем, часть 2 / ЛАНИТ // Хабр [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/companies/lanit/articles/421401>. – Дата доступа: 30.03.2026.
6. Рекомендательные системы: обзор алгоритмов / ProductStar // Хабр [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/companies/productstar/articles/523686>. – Дата доступа: 29.03.2026.
7. Теоретические и практические аспекты построения рекомендательных моделей: типология, архитектура и направления проектирования // КиберЛенинка [Электронный ресурс]. – Режим доступа: <https://cyberleninka.ru/article/n/teoreticheskie-i-prakticheskie-aspekty-postroeniya-rekomendatelnyh-modeley-tipologiya-arhitektura-i-napravleniya-proektirovaniya/viewer>. – Дата доступа: 30.03.2026.