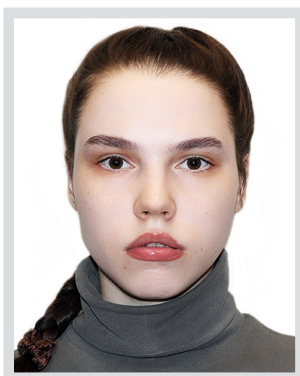


УДК 004.891.2

DOI: 10.24412/1998-5533-2026-2-7-12

© Апон Т.А., 2026

Степень персонализации в рекомендательных системах: сравнительный анализ подходов и критерии выбора

**Апон Татьяна Анатольевна**

Преподаватель кафедры экономики

Национального открытого института г. Санкт-Петербург

(197183, Санкт-Петербург, ул. Сестрорецкая, д. 6; e-mail: tapon-20@ranepa.ru)

В настоящей статье представлен сравнительный анализ подходов к построению рекомендательных систем, в основу классификации которых положена степень персонализации выдаваемого контента. Актуальность темы исследования обусловлена фундаментальным противоречием, неизбежно возникающим на этапе проектирования подобных систем: необходимостью одновременного обеспечения высокой точности индивидуальных предсказаний и сохранения приемлемого уровня вычислительной сложности, интерпретируемости и стоимости внедрения.

Цель работы заключалась в разработке системы критериев и практических инструментов, обеспечивающих осознанный выбор архитектурной стратегии рекомендательной системы на основе ее классификации по степени персонализации. Для достижения поставленной цели решались задачи по выделению и содержательной характеристике четырех классов систем, формированию набора сопоставимых оценочных параметров и построению процедуры принятия проектного решения. Научная значимость исследования состоит в систематизации представлений о градациях персонализации и в обосновании комплекса из семи взаимосвязанных критериев, отражающих ключевой для данной предметной области компромисс между точностью и сложностью. Практическая ценность определяется разработкой блок-схемы и чек-листа – инструментов, применимых на начальных стадиях проектирования и допускающих итеративное уточнение архитектуры по мере развития сервиса.

Основные результаты работы свидетельствуют, что ни один из рассмотренных классов не обладает абсолютным преимуществом. Выбор определяется конкретными условиями проекта, а оптимальной траекторией развития является последовательное движение от простых решений к более сложным по мере накопления данных. Новизна предложенного подхода заключается в переводе проблемы выбора рекомендательной архитектуры из плоскости интуитивных предпочтений в плоскость структурированного сопоставления по явно заданным критериям.

Ключевые слова: рекомендательные системы, степень персонализации, проблема холодного запуска, сравнительный анализ, критерии отбора, коллаборативная фильтрация, гибридные системы

Для цитирования: Апон Т.А. Степень персонализации в рекомендательных системах: сравнительный анализ подходов и критерии выбора // Вестник экономики, права и социологии. 2026. № 2. С. 7–12. DOI: 10.24412/1998-5533-2026-2-7-12.

Рекомендательные системы стали неотъемлемой частью цифрового опыта — от интернет-магазинов до стриминговых сервисов. Они направляют выбор пользователя в условиях информационной перегрузки и непосредственно влияют на ключевые бизнес-метрики: конверсию, средний чек и удержание аудитории. Однако историческое развитие данного класса систем характеризуется фундаментальным противоречием: с одной стороны, наблюдается постоянное усложнение моделей для повышения точности индивидуальных предсказаний, с другой — сохраняется, а в ряде случаев и усиливается, потребность в вычислительно эффективных, интерпретируемых и надежных решениях, основанных на глобальных трендах.

В академической литературе, посвященной рекомендательным системам, обсуждение обычно сводится к сравнению алгоритмов внутри одной парадигмы — например, коллаборативной фильтрации [1; 2]. Однако вопрос о выборе между самими парадигмами на этапе проектирования системы остается концептуально не разработанным. Необходимость персонализации часто принимается как аксиома, без рассмотрения сценариев, в которых ее применение методологически или технически неоправданно. В равной степени отсутствуют четкие критерии для перехода от общих стратегий к персонализированным по мере развития продукта.

Следствием данной неопределенности становятся неоправданные ресурсные затраты: инженерные команды строят сложные *ML*-модели там, где достаточно простых решений, или, напротив, теряют клиентов из-за низкого качества рекомендаций там, где персонализация критична. Данное исследование ставит своей целью восполнить этот пробел, предлагая систему критериев для осознанного выбора стратегии рекомендаций на основе классификации систем по степени персонализации.

Настоящее исследование не претендует на исчерпывающий обзор всех возможных классификаций рекомендательных систем. В фокусе работы находится одна, но основополагающая ось — степень персонализации. Такой выбор обусловлен тем, что именно этот параметр определяет ключевые архитектурные решения на ранних этапах проекта и напрямую связан с компромиссом между точностью и сложностью. В рамках данной классификации выделяют четыре основных подхода.

Под **неперсонализированными** рекомендательными системами понимаются системы, генерирующие единый для всех пользователей набор рекомендаций, основываясь на агрегированных данных о поведении всей аудитории или свойствах самих объектов. К данному классу относятся: ранжирование по популярности (блоки «Хиты продаж», «Популярное»), рекомендации на основе сопутствующих товаров («С этим товаром покупают»), подбор ана-

логов по метаданным («Похожие товары») и анализ ассоциативных правил. Ключевое преимущество таких систем — устойчивость к проблеме «холодного старта», под которой понимается отсутствие данных о новом пользователе или объекте. Именно с неперсонализированных рекомендаций большинство сервисов начинают свою работу, поскольку их реализация не требует знаний о конкретном пользователе и сопряжена с минимальными затратами.

Под **полуперсонализированными** рекомендательными системами понимаются системы, которые не используют детальную историю взаимодействий конкретного пользователя, но при этом отказываются от единой для всех выдачи в пользу сегментирования аудитории. Пользователи объединяются в группы на основе демографических характеристик (возраст, пол), географических и поведенческих признаков (страна, город, время суток) или контекстных сигналов (например, находится ли пользователь в движении, согласно показаниям смартфона) [3]. Типовым подходом к реализации является кластеризация: на основе эмбедингов активных пользователей выделяются кластеры, вычисляются их центроиды, а новые пользователи зачисляются в ближайший кластер и получают топ-*N* наиболее популярных объектов из него [4]. Полуперсонализация выступает как компромиссная стратегия, позволяющая повысить релевантность по сравнению с неперсонализированной выдачей без попытки дорогостоящего предсказания уникальных индивидуальных предпочтений.

Под **персонализированными** рекомендательными системами понимаются системы, формирующие уникальную для каждого пользователя выдачу на основе анализа его персонального профиля — истории взаимодействий, демографических характеристик и/или сходства с другими пользователями. К данному классу относятся три базовых метода. Коллаборативная фильтрация опирается исключительно на матрицу взаимодействий «пользователь-объект» и основана на предположении, что пользователи со схожим поведением в прошлом будут иметь схожие предпочтения в будущем. Контентная фильтрация рекомендует объекты, похожие по своим характеристикам (метаданным) на те, которые пользователь предпочитал ранее. Фильтрация, основанная на знаниях, использует явно заданные правила и экспертные модели предметной области, что делает ее незаменимой в доменах с редкими покупками и сложными техническими характеристиками — недвижимость, автомобили, дорогая бытовая техника [2]. Цель данного класса систем — максимизация индивидуальной релевантности за счет отказа от универсальной модели ранжирования.

Под **гибридными** рекомендательными системами понимаются системы, интегрирующие два и более подхода для компенсации ограничений каждого

из них. В литературе [5] сложилась классификация гибридных систем по двум измерениям: парадигма рекомендаций (какие базовые алгоритмы используются) и дизайн гибридации (как именно эти алгоритмы сочетаются). В рамках последнего измерения выделяют три базовых архитектурных дизайна. Монолитный дизайн предполагает единый алгоритм, использующий признаки из разных источников. Параллельный дизайн подразумевает одновременную работу нескольких независимых рекомендательных с последующим объединением их результатов – смешанным (*mixed*), взвешенным (*weighted*) или с переключением стратегий (*switching*). Конвейерный (последовательный) дизайн предполагает организацию вычислительных процедур в виде цепочки, в которой выходные данные предшествующего алгоритма поступают на вход последующего. Подобная архитектура допускает, в частности, первоначальный отбор объектов с использованием методов, основанных на знаниях, и дальнейшее ранжирование полученного подмножества с применением коллаборативных данных [1]. Таким образом, гибридные системы обеспечивают реализацию сильных сторон различных подходов на соответствующих этапах обработки и формируют условия для постепенного перехода к индивидуализированным моделям повышенной точности.

Для перехода от теоретической классификации к формированию практических рекомендаций автором была разработана система критериев сравнения. В основу ее построения легли два методологических принципа: во-первых, критерии должны отражать фундаментальное противоречие между точностью предсказаний и сложностью их достижения; во-вторых, они обязаны обладать практической значимостью при принятии проектных решений в реальных информационных системах. Руководствуясь данными принципами, автор выделил семь критериев, последовательно раскрываемых ниже.

Точность (релевантность) представляет собой центральную метрику качества рекомендательной системы и характеризует степень соответствия предлагаемых объектов действительным интересам пользователя. Высокая точность напрямую влияет на удовлетворенность и конверсию.

Разнообразие – способность системы выходить за рамки очевидного, предлагая не только самое популярное или строго соответствующее профилю, но и неожиданные, однако интересные позиции. Этот критерий важен для удержания пользователя и предотвращения «пузыря фильтров», когда система замыкает пользователя в узком сегменте контента.

Проблема **холодного старта** – ключевое ограничение персонализированных методов, определяющее, насколько быстро система способна начать приносить пользу новому пользователю или работать с новым объектом. От способа решения этой

проблемы зависит качество первого взаимодействия пользователя с сервисом и, как следствие, его удержание.

Сложность реализации – интегральная оценка затрат на разработку, внедрение и поддержку системы, включающая как требуемые вычислительные ресурсы (мощности серверов, время обучения моделей), так и необходимую квалификацию команды.

Объем требуемых данных – минимальное количество и тип информации, необходимые для того, чтобы подход мог функционировать. Неперсонализированные системы требуют лишь агрегированной статистики, в то время как персонализированные нуждаются в полной истории взаимодействий.

Гибкость – способность системы адаптироваться к изменениям (ассортимента, поведения пользователей, бизнес-приоритетов) без полного перепроектирования. Высокая гибкость позволяет оперативно реагировать на новые требования.

Интерпретируемость – понятность логики рекомендаций как для разработчика (важно для отладки и аудита), так и для конечного пользователя (важно для доверия к системе). Модели типа «черный ящик», такие как глубокие нейронные сети, выигрывая в точности, существенно проигрывают в интерпретируемости.

Совокупность данных критериев позволяет оценивать каждый подход не изолированно по одному параметру, а комплексно, что соответствует реальным задачам проектирования.

Результаты сопоставления четырех подходов по семи выделенным критериям систематизированы в таблице 1.

Представленная таблица 1 демонстрирует, что выбор подхода к рекомендациям не может быть универсальным – он определяется контекстом проекта, доступными ресурсами и требованиями к точности. Обобщая приведенные критерии, автор формулирует следующие рекомендации.

Неперсонализированные рекомендации оптимальны на начальных этапах развития проекта (в качестве MVP), для главных страниц и разделов с общим контентом, а также при работе с неаутентифицированными посетителями. Их главное преимущество – простота и минимальные требования к данным.

Полуперсонализированные рекомендации целесообразно применять для анонимных сессий, региональных и временных подборок, а также в ситуациях, когда данных о пользователе недостаточно для полноценной персонализации, но доступна сегментирующая информация: география, демография, контекст.

Персонализированные рекомендации приобретают приоритетное значение при обслуживании постоянных пользователей и в разделах, где требование максимальной релевантности является критичным, – к таковым относятся ленты формата

Таблица 1

Сравнительная таблица подходов по степени персонализации

Критерий	Неперсонализированные	Полуперсонализированные	Персонализированные	Гибридные
Точность (релевантность)	Низкая: рекомендации одинаковы для всех	Средняя: учитывает контекст и базовые признаки	Высокая: адаптировано под индивидуальные предпочтения	Очень высокая: комбинация методов повышает точность
Разнообразие	Низкое: для всех одинаковые рекомендации, часто повторяются хиты продаж	Среднее: зависит от разнообразия сегментов	Высокое: индивидуальные подборки, находит нишевые интересы	Очень высокое: балансирует популярность и индивидуальность
Проблема холодного старта	Отсутствует: не требует данных о пользователе	Частично решена: использует доступные признаки (сегмент, сессия), но точность ниже, чем у персонализированных	Острая: нужны данные о пользователе для обучения модели	Практически решена: комбинирует персональные и общие данные
Сложность реализации	Низкая: простые SQL-запросы или ручные правила	Средняя: кластеризация, сегментация, базовый анализ сессий	Высокая: требуются сложные ML-алгоритмы	Очень высокая: интеграция нескольких алгоритмов, настройка весов
Объем требуемых данных	Минимальный: агрегированная статистика	Средний: данные для сегментации	Большой: полная история взаимодействий пользователя и доступ к профилю	Очень большой: все типы данных + метаданные о товарах/контенте
Гибкость	Высокая: система основана на простых правилах и статических списках, которые легко менять вручную	Средняя: требует обновления правил сегментации или параметров контекста	Низкая: модель требует переобучения при изменении ассортимента или паттернов поведения	Средняя: возможно отключать/добавлять компоненты и настраивать их вес, но требуется координация между блоками
Интерпретируемость	Высокая: очевидно, почему рекомендован товар (популярное, новинки)	Средняя: объяснима на уровне сегментов (география, демография), но не персонализирована	Низкая: модели типа «черный ящик»	Средняя/низкая: зависит от комбинации методов

Источник: составлено автором.

«Для вас» и персональные подборки. Вместе с тем внедрение данного класса систем сопряжено с необходимостью накопления значительных массивов пользовательских данных и осуществления инвестиций в ML-инфраструктуру.

Гибридные системы, в свою очередь, представляют собой решение, сочетающее наибольшую гибкость с максимальной сложностью реализации. Их применение становится оправданным на крупных платформах – маркетплейсах, стриминговых сервисах, – характеризующихся, с одной стороны, доступностью всех типов данных, а с другой – высокой ценой ошибки рекомендательного механизма. Гибридный подход дает возможность смягчить проявления проблемы холодного старта и выстроить плавную траекторию перехода от общих стратегий ранжирования к индивидуальным моделям по мере накопления сведений о пользователе.

В целях практического применения предложенной системы критериев автором разработаны два инструмента поддержки принятия решений: блок-схема (рис. 1) и чек-лист (рис. 2). Блок-схема реализует логику пошагового продвижения от исходных условий проекта к обоснованному выбору архитек-

турного подхода. Чек-лист, в свою очередь, представляет собой структурированный набор условий, позволяющих оперативно оценить степень применимости каждой из стратегий к конкретному проекту.

Указанные инструменты могут применяться как независимо друг от друга, так и совместно. Существенно подчеркнуть, что они не предназначены для выработки единственно верного решения – их функция заключается в структурировании пространства альтернатив и допущении итеративности проектного процесса. На практике это означает, что проект может стартовать с внедрения неперсонализированных рекомендаций, пройти через стадию полуперсонализации по мере накопления данных и впоследствии эволюционировать к персонализированным либо гибридным архитектурам при достижении аудиторией и инфраструктурой требуемых характеристик.

Выводы

Проведенное исследование позволяет сформулировать три основных вывода:

Во-первых, не существует универсально лучшего подхода к построению рекомендательной системы. Выбор всегда определяется конкретным контекстом проекта: доступными данными, ресурсами на

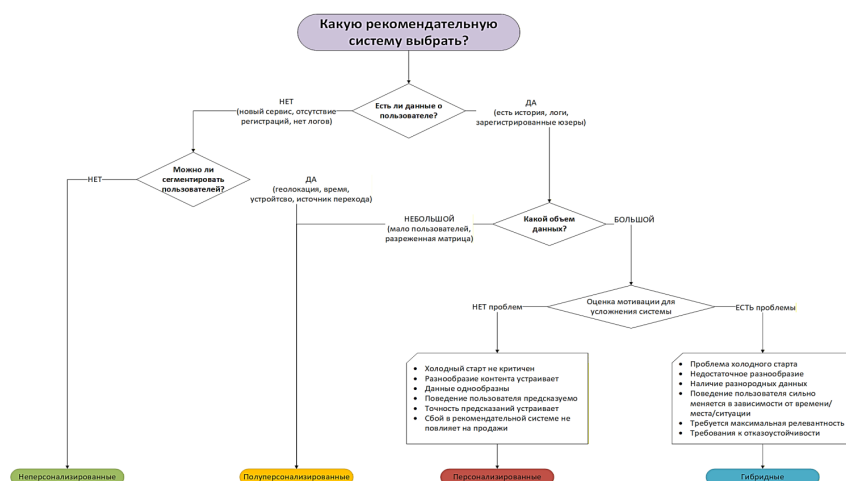


Рис. 1. Блок-схема
«Какую рекомендательную систему выбрать?»

Какую стратегию рекомендаций выбрать?

Принятие решения: ✓ от 3 и более → стратегия применима

НЕПЕРСОНАЛИЗИРОВАННЫЕ

- ☐ Данных о пользователях нет / почти нет
- ☐ Нужно запустить MVP и проверить гипотезу
- ☐ Скорость реализации критична
- ☐ Бюджет на разработку ограничен
- ☐ В команде нет ML-инженеров

ПОЛУПЕРСОНАЛИЗИРОВАННЫЕ

- ☐ Есть данные о геолокации / времени суток / сезоне / типе устройства / демографические данные
- ☐ Пользователям из разных мест / в разное время нужны разные товары
- ☐ Данных на полноценную персонализацию пока мало
- ☐ Есть аналитик, который может настроить сегменты

ПЕРСОНАЛИЗИРОВАННЫЕ

- ☐ Есть история действий пользователей (лайки, отзывы, клики, просмотры, покупки)
- ☐ Есть регистрация / авторизация пользователей
- ☐ Уже есть достаточно данных (тысячи пользователей, десятки тысяч действий)
- ☐ Нам важно удерживать постоянных клиентов
- ☐ Бюджет позволяет нанять ML-команду

ГИБРИДНЫЕ

- ☐ Высокая конкуренция в нише
- ☐ Ошибка в рекомендациях критична, можно потерять клиента
- ☐ Доступны все типы данных (поведенческие + метаданные + контекст)
- ☐ Есть ресурсы на сложную архитектуру (команда, бюджет, сервера)
- ☐ Нужно одновременно работать и с новыми, и с активными пользователями
- ☐ Требуется максимальная точность (хотим быть лучшими в своём деле)

Рис. 2. Чек-лист «Какую стратегию рекомендаций выбрать?»

разработку, требованиями к интерпретируемости и стадией жизненного цикла продукта. Каждый из четырех рассмотренных классов занимает свою нишу и оптимален в определенных условиях;

Во-вторых, выбор стратегии – это всегда компромисс между точностью, сложностью, объемом данных, интерпретируемостью и другими факторами. Предложенная система из семи критериев позволя-

ет принимать этот компромисс осознанно, а не интуитивно;

В-третьих, лучшей практикой является начало с простого и эволюционное усложнение системы по мере роста проекта. MVP на неперсонализированных рекомендациях – это нормально. Полуперсонализация становится целесообразной, когда накоплено достаточно данных для выделения пользовательских кластеров. Персонализация и гибридные архитектуры оправданы тогда, когда база выросла, данные накоплены, и точность рекомендаций действительно начинает влиять на ключевые показатели продукта.

Литература:

- Jannach D., Zanker M., Felfernig A., Friedrich G. Recommender Systems. Cambridge: Cambridge University Press, 2010. URL: <https://www.researchgate.net/publication/235910467>.
- Ricci F., Rokach L., Shapira B., Kantor P.B. Recommender Systems Handbook. New York: Springer, 2015. URL: https://www.researchgate.net/publication/227268858_Recommender_Systems_Handbook.
- Меньшикова Н.В., Портнов И.В., Николаев И.Е. Обзор рекомендательных систем и возможностей учета контекста при формировании индивидуальных рекомендаций // Academy. 2016. № 6 (9). С. 20–22.
- Briand L., Salha-Galvan G., Bendada W., Morlon M., Tran V. A Semi-Personalized System for User Cold Start Recommendation on Music Streaming Apps // arXiv.org. 2021. 7 June. URL: <https://arxiv.org/abs/2106.03819>.
- Burke R. Hybrid Recommender Systems: Survey and Experiments // User Modeling and User-Adapted Interaction. 2002. Vol. 12. № 4. P. 331–370. <https://doi.org/10.1023/A:1021240730564>.
- Фальк К. Рекомендательные системы на практике / Пер. Д.М. Павлов. М.: ДМК Пресс, 2020. 448 с.
- Чжан Л. Анализ компромисса между точностью и сложностью идентифицированных моделей динамических систем // Системы анализа и обработки данных. 2024. № 2 (94). С. 85–93.
- Чжан Чжибо, Афанасьев Г.И. Основные технологии и перспективы эволюции персонализированных рекомендательных систем // E-Scio. 2022. № 4 (67). С. 309–320.

Degree of Personalization in Recommender Systems: a Comparative Analysis of Approaches and Selection Criteria

Tatyana Anatolyevna Apon

National Open Institute, St. Petersburg (Saint Petersburg, Russia)

This article presents a comparative analysis of approaches to building recommender systems based on the degree of personalization of the content they provide. The relevance of this research topic stems from a fundamental contradiction that inevitably arises during the design phase of such systems: the need to simultaneously ensure high accuracy of individual predictions while maintaining an acceptable level of computational complexity, interpretability, and implementation cost.

The goal of this study was to develop a system of criteria and practical tools to support an informed choice of an architectural strategy for a recommender system based on its classification by degree of personalization. To achieve this goal, we identified and substantively characterized four classes of systems, developed a set of comparable evaluation parameters, and developed a design decision-making procedure.

The scientific significance of this study lies in its systematization of concepts regarding gradations of personalization and the substantiation of a set of seven interrelated criteria reflecting the key tradeoff between accuracy and complexity in this subject area. The practical value is determined by the development of a flowchart and checklist-tools applicable in the initial design stages and allowing for iterative refinement of the architecture as the service evolves.

The main results of the work indicate that none of the considered classes has an absolute advantage. The choice is determined by the specific project conditions, and the optimal development trajectory is a consistent progression from simple solutions to more complex ones as data accumulates. The novelty of the proposed approach lies in its shift from the problem of choosing a recommendation architecture to one of structured comparison based on explicit criteria.

Keywords: recommender systems, degree of personalization, cold start problem, comparative analysis, selection criteria, collaborative filtering, hybrid systems

