**Документация, содержащая описание функциональных характеристик модуля программного обеспечения**

**Модуль «Рекомендательная система для стриминговых сервисов»**

**г. Москва 2025**

**Содержание**

[**1. Введение 3**](#_heading=h.cmm46i126d55)

[**2. Назначение модуля 3**](#_heading=h.ts3wpoedj4l9)

[**3. Функциональные возможности 3**](#_heading=h.ev1x8s185mjq)

[**4. Архитектура и принцип работы 4**](#_heading=h.mzas2gsn8rlr)

[**5. Используемые языки программирования и фреймворки 5**](#_heading=h.c36rm2r1e9mu)

## **1. Введение**

Рекомендательная система для стриминговых сервисов представляет собой инструмент, использующий алгоритмы анализа данных и машинного обучения для предоставления пользователям персонализированных рекомендаций контента.

Основная цель системы – повышение вовлеченности пользователей, увеличение времени просмотра и улучшение пользовательского опыта.

Система учитывает данные, включая историю просмотров, лайки, комментарии, подписки, тренды и предпочтения схожих пользователей, чтобы формировать наиболее релевантные рекомендации.

Данный модуль интегрируется в платформу и обеспечивает рекомендации для различных типов пользователей.

## **2. Назначение модуля**

Модуль "Рекомендательная система для стриминговых сервисов" предназначен для персонализированного подбора видеоконтента на основе анализа данных о поведении пользователей. Основные задачи модуля:

* Удержание пользователей за счет релевантных рекомендаций.
* Персонализация контента на основе индивидуальных предпочтений.
* Открытие нового контента для пользователей.
* Увеличение вовлеченности за счет подбора контента, стимулирующего активное взаимодействие (лайки, комментарии, подписки).

## **3. Функциональные возможности**

### Основные возможности модуля:

1. Персонализированные рекомендации – подбор контента на основе истории просмотров, лайков, комментариев, поисковых запросов.
2. Рекомендации по категориям – учет жанра и тематической направленности контента.
3. Контекстные рекомендации – подбор похожих видео, формирование плейлистов.
4. Коллаборативные рекомендации – подбор контента, популярного среди пользователей с аналогичными интересами.
5. Алгоритмические рекомендации – использование машинного обучения для более точных рекомендаций.
6. Рекомендации для новых пользователей – формирование рекомендаций при отсутствии данных (решение проблемы "холодного старта").
7. Трендовые рекомендации – анализ актуальных тем и трендов.
8. Мультиплатформенные рекомендации – синхронизация данных между устройствами (смартфоны, ПК, Smart TV).
9. Экспериментальные рекомендации – тестирование новых алгоритмов, A/B тестирование.
10. Учет пользовательского контента (UG**C)** – включение пользовательского контента в рекомендации.

## **4. Архитектура и принцип работы**



Рис. 1 диаграмма концептуальной архитектуры модуля рекомендательной системы для стриминговых сервисов

Модуль рекомендательной системы построен на многослойной архитектуре, включающей следующие уровни:

**Сбор данных**

Система анализирует пользовательские данные, такие как:

Просмотренные видео.

Реакции (лайки, дизлайки, комментарии).

Поисковые запросы.

Подписки на каналы.

**Анализ и обработка данных**

Используются следующие методы и алгоритмы:

Коллаборативная фильтрация – анализ поведения схожих пользователей.

Контентная фильтрация – анализ схожих характеристик контента (жанр, теги, описание).

Гибридные методы – комбинация коллаборативной и контентной фильтрации.

Машинное обучение – использование нейросетей и моделей для улучшения рекомендаций.

**Генерация рекомендаций**

На основе анализа данных система формирует персонализированные подборки контента, включая:

Видео, схожие с уже просмотренными.

Популярные видео среди пользователей с аналогичными интересами.

Новые видео из подписок.

Трендовые видео.

**Интерактивная адаптация**

Система анализирует реакцию пользователей на предложенные рекомендации (например, пропуск видео, лайки/дизлайки) и адаптирует последующие подборки в соответствии с их предпочтениями.

## **5. Используемые языки программирования и фреймворки**

**Python** – основной язык программирования.

**Модуль реализован с использованием современных технологий и инструментов:**

**Серверная инфраструктура**

Kubernetes – оркестрация контейнеров.

Docker – контейнеризация сервисов.

Nginx – веб-сервер и балансировщик нагрузки.

FastAPI + Uvicorn – фреймворк для создания API.

**Базы данных**

PostgreSQL – основная база данных.

Redis – кэширование запросов.

OpenSearch – полнотекстовый поиск.

ClickHouse – хранение больших объемов логов.

**Машинное обучение и анализ данных**

Библиотеки:

* + scikit-learn, PyTorch, XGBoost, CatBoost – обучение моделей.
	+ Sentence Transformers – работа с текстовыми эмбеддингами.
	+ NumPy, SciPy, Pandas, Polars – анализ данных.
	+ Dagster, Apache Spark – обработка данных.
	+ Optuna – гиперпараметрическая оптимизация моделей.

**Логирование и мониторинг**

Prometheus + Grafana – сбор и анализ метрик.

Grafana Loki – логирование событий.

**Интеграционные сервисы**

API Rutube – взаимодействие с видеоплатформой.

Яндекс.Метрика, AppMetrica – анализ пользовательского поведения.

Kafka – потоковая обработка данных.