

Ссылка на  
репозиторий:<https://github.com/NastyaKha/-/blob/main/README.md>



# Разработка системы рекомендации друзей в социальной сети

Проект выполнили: Ха Анастасия  
Шамара Софья  
Шустиков Андрей  
Прунь Егор

Ссылка на  
репозиторий: <https://github.com/NastyaKha/-/blob/main/README.md>

# Актуальность создания рекомендательной системы друзей

Социальная сеть Вконтакте очень популярна и в ней зарегистрировано большое количество людей. Это большая площадка для общения, но самая главная проблема – это подобрать контингент для общения. Хотя платформа уже имеет рекомендательную систему, но она не совершенна. Основная её настройка искать друзей по общим знакомым или включать фильтрацию по общим признакам. Но это не всегда помогает для нахождения друзей или знакомых.

## Что, если:

- ❓ ваш знакомый не учился в одной школе с вами и вы не знаете название
- ❓ вы не знаете где именно он живет.
- ❓ вне помните его фамилию или может он её сменил.
- ❓ не знаете место работы или он ещё не работает.
- ❓ нет общих знакомых.

Тогда будет просто не возможно найти нужного человека. Но зато вы часами разговаривали о любимой музыке, фильмах и интересах





## Цель работы

Выявить недостатки рекомендательной системы друзей в социальной сети Вконтакте;

Придумать более оптимальный алгоритм для нахождения знакомых и новых знакомств.

## Задачи

Рассмотреть принцип работы существующей системы;

Выявить ситуации, которые не рассматривает система;

Составить алгоритм фильтрации и рассмотреть наиболее оптимальный тип рекомендательной системы для отбора



# Анализ области

**Система рекомендаций друзей** – это такая система, которая автоматически предлагает потенциальных друзей пользователям

Эту систему используют различные социальные сети: Facebook, Instagram, LinkedIn, Вконтакте, Одноклассники и т.д .

Система рекомендаций – это система, которая используется для рекомендации ресурсов, которые могут быть интересны пользователю, исходя из интересов и / или предпочтений пользователей. В социальных сетях существуют разные типы рекомендательных систем, и каждая из этих систем использует разные технологии для рекомендации





# Виды рекомендательных систем

## Фильтрация на основе содержимого (content-based filtering)

На основе содержимого пытается предсказать знакомы ли пользователи, используя его историю посещения сообществ, страниц других пользователей, общих знакомых и поставленных лайков на фото и видео.

## Коллаборативная фильтрация (collaborative filtering)

User-Based Collaborative Filtering (UBCF): Рекомендации строятся на основе схожести между пользователями. Если два пользователя имеют похожие предпочтения, то они могут получать рекомендации на основе того, что понравилось другому пользователю. Item-Based Collaborative Filtering (IBCF): Рекомендации строятся на основе схожести между товарами или контентом. Если пользователь предпочитал определенный товар, то ему будут рекомендованы похожие товары, которые другие пользователи с похожими интересами также предпочли.

## Фильтрация, основанная на знаниях (knowledge-based filtering)

В этом случае можно добавить фильтры: пол, любимый жанр фильмов, музыки, игры, сообщества. После этого рекомендательная система подбирает наиболее подходящих людей из каталога.

## Гибридные рекомендательные системы (hybrid filtering)

Гибридные рекомендательные системы сочетают разные подходы, то есть одновременно включаются механизмы фильтрации, основанной на контенте, и коллаборативной фильтрации. Концепция данного подхода заключается в том, что разные системы могут эффективно взаимодействовать, тем самым компенсируя недостатки друг друга.

# Признаки для рекомендаций

В социальной сети Вконтакте уже есть определенная фильтрация для нахождения друзей. Это возраст, город, место учебы, место работы, армия и наличие хотя бы одного общего друга.

Предлагаем добавить туда:

- поиск по группам в Вконтакте,
- интересы по видео (можно совпадение по интересам видеороликов)
- совпадение по музыке (фильтр по совпадению похожей музыки существует, но добавить на прямую человека пока нет возможности)
- поиск по играм на сайте (для тех кто ищет людей для создания команды в игре)

Все представленные признаки есть в разделах социальной сети, так что отдельно просить пользователя вносить данные не придется.

Необходимо проанализировать выбор пользователей в каждом сегменте и при совместимости близкого образа жизни предлагать знакомство.

Либо если это общая группа или игра предложить этим людям стать друзьями.

Можно добавить кнопки для расширенного поиска:

- Поиск по группам
- Поиск по названию игры

А остальные признаки фильтровать по запросам на сервисах и их схожестью. И добавить в ленту, сделав для них специальные обозначения, по какому признаку они похожи

Рейтинг в ленте можно тоже поставить на фильтр.

Наверх рейтинга ставить людей, которые:

- Периодически входят в Вконтакт. (чтобы избежать заброшенных аккаунтов)
- У которых соотношение добавления в друзья и отписки хотя бы в 70% разницы (для избегания рекламных ботов и странных личностей)

Можно придумать какие-нибудь задания для прохождения их с новыми друзьями и получения за них стикеров.

Создать архив анкет или историю анкет, от которых отказались, для возможно вернуться к ним. (для vip пользователей)



# Формулировка критериев

**Основная идея работы** – создание рекомендательной системы для социальной сети Вконтакте для нахождения знакомых людей или интересных для нового знакомства.

Также предполагалось, что для пользователя может иметь значение по какому критерию искать человека, то есть **вторая** из главных целей – вписать как можно больше признаков отбора для алгоритмов рекомендаций и возможность выбирать по определенному признаку

**Третий критерий** – проверка страниц пользователей на частоту пользования страницей. Тем самым увеличить шанс общения реальных пользователей.

**Четвертый критерий** – это возможность простота пользования фильтрацией.





# Сравнительный анализ

На основе заданных критериев для создания системы рассматривать можно два варианта алгоритмов. Наиболее популярный – это **коллаборативная фильтрация** и **гибридный способ**.

Если использовать **коллаборативную фильтрацию**, то трудности возникнут с составлением матрицы так, как нужно будет учитывать много различных факторов для фильтрации, с учетом очень большого количества пользователей. Возможно лучше использовать модель нейронной сети.

Проблема будет с холодным стартом для тех, кто только зарегистрировался и ещё особо никуда не заходил и мало информации о себе предоставил.

Неправдивая информация в сведениях, при заполнении анкеты.

## Гибридная фильтрация

В системе гибридных рекомендаций используются методы глубокого обучения. Система может быть более точная и находить больше сходств сама. Очень удобна для такого количества данных и неоднозначных запросов. Но для создания такой системы потребуется много времени и возможны проблемы с закрытой информацией. Чем выше уровень конфиденциальности, тем ниже точность рекомендации.

Чтобы компенсировать проблему низкой точности рекомендаций, вызванную технологией глубокого обучения, мы можем объединить технологию глубокого обучения с технологией обучения с подкреплением синтезировать преимущества двух технологий машинного обучения и использовать технологию обучения с глубоким подкреплением для дальнейшего повышения точности системы рекомендаций.



# Виды используемых методов

## Метод 1

Основанный на образе жизни,  
объединяющем людей



Первичные группы  
(члены семьи)



Вторичные группы  
(рабочий коллектив)



Неформальные группы  
(друзья)



Малые группы  
(школьный класс)



Большие группы  
(нация, сословие)



Формальные группы  
(спортивная команда)

## Метод 2

На основе реального местоположения и  
интересов пользователя

