

HACKATHON

# Рекомендательная Система



Las Teteras Desesperadas

# Цель проекта



Разработка передовой ML-рекомендательной  
системы для персонализированного подбора  
образовательных программ и ВУЗов



Система упрощает процесс выбора образовательных учреждений для абитуриентов, предоставляя индивидуальные рекомендации на основе их уникальных интересов и предпочтений

— LAS TETERAS DESESPERADAS



## Парсинг данных о пользователях

- проверка столбцов
- загрузка

## Формирование датасета

## Очистка интересов от шумов

## Предобработка интересов

## Фильтрация вспомогательных таблиц

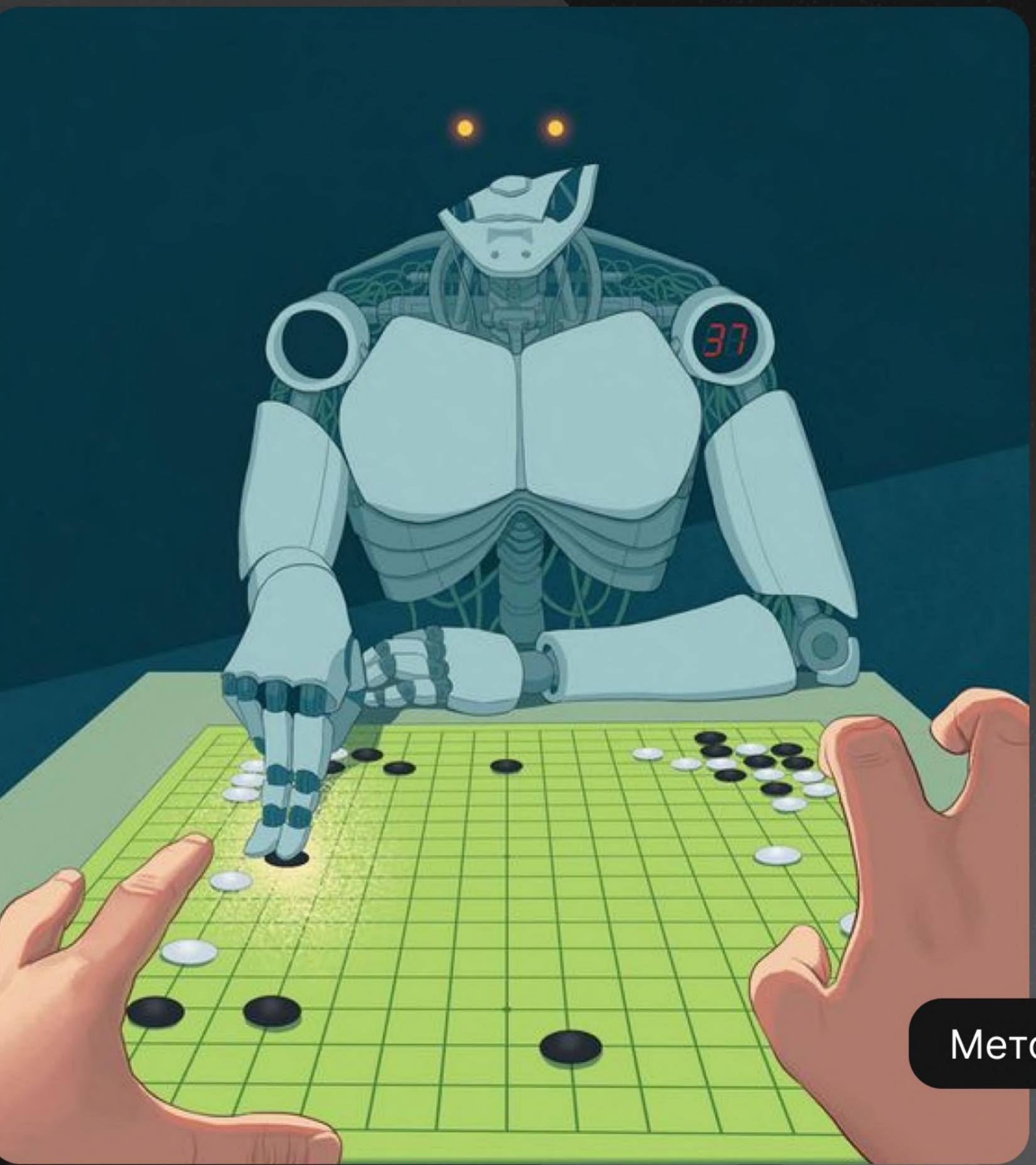
## Модель рекомендации по интересам



# Алгоритм рекомендаций

В системе используется алгоритм  
рекомендаций на основе косинусного  
сходства.

Для преобразования строк интересов  
пользователей используется  
разреженная матрица TF-IDF.



parser.py:

Загрузка данных из API-сервиса, парсинг данных, сохранение данных в CSV-файлы.

---

int\_checker.py:

Проверка данных на наличие строк с некорректными типами данных, удаление таких строк.

---

loader.py:

Очистка данных от дубликатов, удаление строк с неполными данными, объединение данных в единый датафрейм, создание вспомогательных таблиц.

---

aux\_filter.py:

Очистка вспомогательных таблиц от дубликатов, удаление строк с неполными данными.

---

pre\_filter.py:

Удаление строк с пустыми интересами, удаление рекламных слов из интересов.

---

noise\_filter.py:

Удаление стоп-слов из интересов, удаление дубликатов из интересов.

---

preprocessor.py:

Преобразование интересов в векторы TF-IDF.

---

recommender.py:

Обучение модели рекомендаций, выдача рекомендаций.

---



# Система может использоваться для



## Подбор ВУЗов и факультетов

на основе схожести интересов  
пользователя и других пользователей,  
уже выбравших ВУзы



## Подбор ВУЗов

расположенных в конкретном регионе

# Перспективы развития

- Расширение географии ВУЗов
- Подключение иностранных языков
- Включение средних учебных заведений



# Над проектом работали



Загальский Игорь



Санникова Юлия



Мухачев Иван



Антропова Наталия



Майнгерт Владимир



Аванесян Тачат

# UniGoTo

Рекомендательный сервис ВУЗов и специальностей на основе технологии Искусственного

Интелле

 <https://unigoto.streamlit.app/>

Пропустить рекомендации по региону

Введите ваши данные и предпочтения:

О себе

Умный, оптимист

Сфера

Информатика, программирование

Любимые книги

Макроэкономика

Любимые игры

Тетрис, Цивилизация

интересы и хобби

Путешествия, развлечения

Ваш город

СКВА

Справить

данные:

умный, оптимист инфо

Ваш город: М

# Спасибо за внимание



# Las Teteras Desesperadas

Рекомендательная система использует разреженную матрицу TF-IDF для преобразования строк интересов пользователей и алгоритм косинусного сходства для вычисления схожести их интересов.

ТОП-20 рекомендованных ВУЗов\*

ВУЗ

КГАСУ

2) ВГНА

3) МГАКХИС (МФ)

4) ЗФ СГА

5) НГСУ (бывш. УПС)

6) МИУ (бывш. АГУ, АГ

7) КФ СГА

8) НГТУ (НЭТИ)

9) ВлГУ им. А. Г. и Н. Г. Столетовых

10) СПбГУ

11) МОИУ (бывш. ИБПИТ)

12) МГУ им. Н. П. Огарева

13) СГЭУ

14) КФ АТИСО

15) СГАУ им. С. П. Королева

16) РАНХиГС при Президенте РФ (бывш. Университет

18) Университет управления (бывш.

20) СПбГУ

Факультет

Общей архитектурно-художественной подготовки

Факультет информационных технологий

Факультет экономики и производственного менеджмента

Экономика

Автоматика, телемеханика и связь

Факультет экономики и управления (бывш. Институт менеджмента, Институт информационных технологий)

Юриспруденция

Институт дистанционного обучения

Юридический институт

Биологический факультет

Менеджмента и информационных технологий

Факультет математики и информационных технологий



Сходство

0.2992

0.2947

0.284

0.2747

0.267

0.2606

0.2559

0.2501

0.2488

0.2487

0.2452

0.2416

# Las Teteras Desesperadas

## SKILLFACTORY



Уральский  
федеральный  
университет

имени первого Президента  
России Б.Н. Ельцина