

ПРИКЛАДНАЯ ПРОЕКТНАЯ РАБОТА ПО ТЕМЕ
КЛАССИФИКАЦИЯ ВОЗРАСТА ПО ТРАНЗАКЦИЯМ ПОЛЬЗОВАТЕЛЯ

Участники :

ФИО: Жимоедов Денис Евгеньевич

Регион: Красноярский край

Город: г. Красноярск

Образовательная организация: МАОУ СШ №7 с УИОП

Личный вклад: создание искусственного интеллекта,
определяющий возрастную группу пользователя по транзакциям,
сбор данных для данной задачи.

ФИО: Чмилева Полина Дмитриевна

Регион: Воронежская область

Город: г. Воронеж

Образовательная организация: МБОУ СОШ №62

Личный вклад: Описание проблемы, актуальности, цели и задач,
разбор аналогов проекта и особенностей нашего проекта,
способов привлечения ресурсов, потенциальных заказчиков.

ПРОБЛЕМА

Доступные данные о транзакциях не используются банками для продвижения рекламы, из-за отсутствия программы, определяющей возраст клиента по его покупкам.

ЦЕЛЬ

Создать искусственный интеллект, который будет определять возрастную группу потребителя, основываясь на его транзакции

АКТУАЛЬНОСТЬ

В настоящее время у банков нет возможности подбирать нужную и полезную для пользователя рекламу, из-за незнания к какой возрастной категории относится клиент, выдавая продукты, выбор которых происходит по системе запросов в интернете, что предоставляет дискомфорт пользователю.

ЗАДАЧИ

1. Изучить рынок
2. Получить данные пользователя про транзакциям пользователя и оценки его возрастов
3. Написать код для ИИ
4. Проверить качество ИИ

ПОТЕНЦИАЛЬНЫЕ ЗАКАЗЧИКИ

Банки, которые хотят эффективного сотрудничества с магазинами товаров и услуг (продавать нужный продукт потребителю). Они будут предлагать скидки для магазинов пользующихся этим банком, а банк будет получать клиентскую базу.

ПОТЕНЦИАЛЬНЫЕ ПОТРЕБИТЕЛИ

Магазины, которые хотят использовать эффективную рекламу своего продукта или услуги.

ПОТЕНЦИАЛЬНЫЕ ПОЛЬЗОВАТЕЛИ

Пользователи, которые будут видеть рекламу в банковской базе, которая им нужна. Вследствии чего они будут ей пользоваться.

ПЛАН РАБОТЫ

- 1) Найти и скачать данные для ИИ
- 2) Сделать ИИ
- 3) Использовать ИИ на тестовых данных
- 4) Узнать качество ИИ

РЕСУРСНОЕ ОБЕСПЕЧЕНИЕ

Компьютер, данные от “Академия искусственного интеллекта для школьников” при поддержке Сбербанка, [Jupyter Notebook](#).

СПОСОБЫ ПРИВЛЕЧЕНИЯ

Чтобы увеличить качество ИИ нужно больше данных для ее обучение, для этого договоримся с нашими заказчиками(банки), которые будут предоставлять нам id пользователей, информацию о их транзакциях и их примерную возрастную группу (каждая полученная информация будет анонимизированна).

ХОД РАБОТЫ

Сбор данных:

Данные для нашего проекта предоставила “Академия искусственного интеллекта для школьников” при поддержке Сбербанка ([файл с данными](#)). Эти данные разделены на пять групп:

- 1) группа транзакций(`small_group_description`): номер группы и класс данной группы.
- 2) тренировочная(`transactions_train`) : уникальный номер клиента, дата транзакции, группа транзакции, сумма транзакции.
- 3) ответ на тренировочную группу(`train_target`) : уникальный номер клиента и его возрастная группа.
- 4) тестовая(`transactions_test`) : уникальный номер клиента, дата транзакции, группа транзакции, сумма транзакции.
- 5) ответ на тестовую группу (`test`): уникальный номер клиента.

Использование данных:

Считаем агрегационные признаки (сумма, стандартное отклонение, медианное значение, среднее арифметическое, минимум, максимум, квантили и количество транзакций) для каждого из клиентов и для каждого клиента по всем группам транзакций. Обучим нейросеть с оптимально найденными настройками([где нашли](#)). Используем нашу систему на тестовых данных и получим ответы.

Оценим эффективность:

Сравним данные, которые мы должны были получить и получим точность правильных ответов. В итоге у нас 64,61% правильных по тестам “Академии искусственного интеллекта для школьников” .

[Ссылка на код программы](#) представленная в виде Jupyter Notebook

Мы получили программу, которая будет определять , достаточно часто, возрастную группу клиента.

```
In [39]: import time
import os

current_timestamp = int(time.time())
submission_path = 'submissions/{}.csv'.format(current_timestamp)

if not os.path.exists('submissions'):
    os.makedirs('submissions')

print(submission_path)
submission.to_csv(submission_path, index=True)

submissions/1575149049.csv
```

На тестовых значениях такое решение предугадало 64,61%

АНАЛОГИ

Самым эффективным методом рекламы на сегодня - таргетированная реклама. Она получает информацию о поиске и переходах по ссылке. В результате, этот искусственный интеллект определяет товар, который искал пользователь. И находит разновидности данного товара и выдает его пользователю.



НАШ МЕТОД



Мы получаем информацию о транзакциях пользователя. Искусственный интеллект по этим данным переводит ее в возрастную группу и отправляет рекламодателю, которому нужна данная возрастная группа. В результате чего, реклама под которую входит данная возрастная группа высвечивается пользователю.

СРАВНИТЕЛЬНАЯ ТАБЛИЦА

	таргетированная реклама	Реклама с нашим искусственным интеллектом
Показывает желаемый для пользователя продукт	чаще нет (минус)	чаще да (плюс)
Настойчивая реклама	да (минус)	нет (плюс)
База клиентов	для пользователей интернета (плюс)	для пользователей банка(минус)

Таргетированная реклама показывает поиск основанный на клике и поиске. Пользователь, не всегда, ищет то или переходит на сайты, что он хочет приобрести или что попадает в его категорию интереса. В нашем случае реклама будет попадать в категорию, почти однозначно, потому что она основывается на реальных покупках пользователя.

Таргетированная реклама выскакивает при посещении, почти любого сайта в интернете, это часто пугает пользователей и они не переходят по ней. В нашем случае, она может высвечиваться на экране банкоматов нашего банка и в мобильном приложении.

База пользователей, которые увидят рекламу таргетированную больше, чем нашу (т.к. пользователей интернета больше, чем у всех банков вместе взятых)

СПИСОК ЛИТЕРАТУРЫ

- 1) [Все данные используемые для создания ИИ в проекте](#)
- 2) [Настройка классификатора](#)
- 3) [Про таргетированную рекламу](#)