

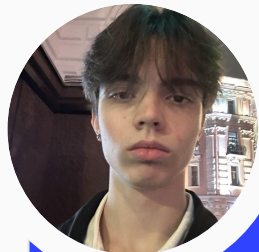
Решение

**“ИИ и цифровые
технологии в
социологических
исследованиях”**

Команда “aCUtone!”



Команда aCUTone!



Тимур Небольсин

- Суперфиналист кейс-чемпионата MOS.MШУ
- Закончил 5 лет хип-хопа
- Ходит только в очках



Михаил Гаврилов

- Призер олимпиады Т-Банка по математике
- Окончил музыкальный колледж
- Участник сборной ЦУ по Kaggle



Дарья Ботягина

- Финалистка DANO и НТО АБП
- Месяц стажировалась в КРОК
- Без ума от грызунков, особенно хомячков



Егор Стариков

- Победитель хакатона ЦУ x Т-Банк
- Фанат дата-аналитики и красивых графиков
- Неиссякаемый источник хорошего настроения в команде



Иван Курбан

- Победитель в хакатоне от Яндекс
- Стипендиат Т-Банка
- Участник сборной ЦУ по Kaggle

Roadmap обучения нейросети



Загрузка данных

- Создание единого датасета со всеми размеченными данными

Предобработка

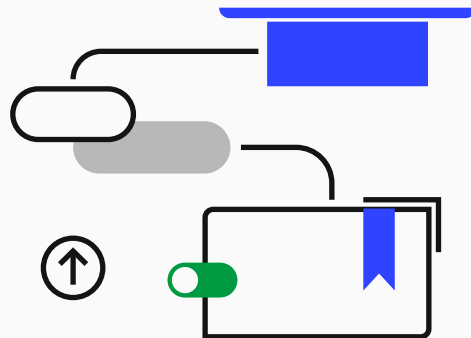
- Очистка текста от посторонних знаков
- Удаление стоп-слов
- Стемминг

Обучение модели

- Добавление **ключевых слов** по каждой категории
- Разделение размеченных данных на train & test
- Векторизация
- Подбор гиперпараметров

Предсказание на основе модели

- Получение итоговых данных
- Были использованы библиотеки Sklearn, NLTK, Pandas, Numpy



Результаты



83%

Ассурасу, полученная с применением кросс-валидации

471

В категории Учебная и внеучебная
вовлеченность

427

В категории Социально-бытовые условия

190

В категории Финансовые условия

320

В категории Лояльность к ВУЗу

Практическая значимость



Наша нейросеть научилась категоризировать информацию самостоятельно. **Как мы это используем?**

Снижение затрат

Затраты на использование нейросетевых технологий ниже, чем расходы на оплату труда сотрудников, особенно при обработке больших объемов данных.

Увеличение скорости обработки

Нейросеть осуществляет анализ тысяч постов за секунды, что несравнимо больше скорости работы человека

Постоянная доступность

Нейросеть функционирует круглосуточно, без выходных, отпусков и перерывов.

Универсальность

Нейросеть можно настроить для работы с разными типами данных, например: тексты, отзывы, анкеты и пр.

Почему именно ИИ?

На примере форума*

На форумах появляются миллионы тредов, наилучший ответ всегда подтверждается с помощью человека, но с помощью функции нашей ИИ, мы можем верифицировать ответ, не затрачивая человеческих сил.

Очевидно, что такой подход будет оптимизированнее и доходнее, чем человеческий труд.

Пример категоризации ответов



Верифицированный
Хороший
Неполный
Выброс



Как возвести 2 в квадрат?



Умножь двойку на саму себя



Представь 2, взятое по два



Легко! $2 \times 2 = 4$



Арифметика — раздел математики, изучающий числа, их отношения и свойства

