



# Кейс VK

Маруся не отвечает на реплики из телевизора

цифровой  
прорыв —↑

сезон: III

Хакатон СКФО



GrowAI

# ПОСТАНОВКА ЗАДАЧИ

- 01 Обучение модели, отличающей голосовые команды пользователя от внешнего шума
- 02 Задача бинарной классификации
- 03 Метрика - ROC-AUC



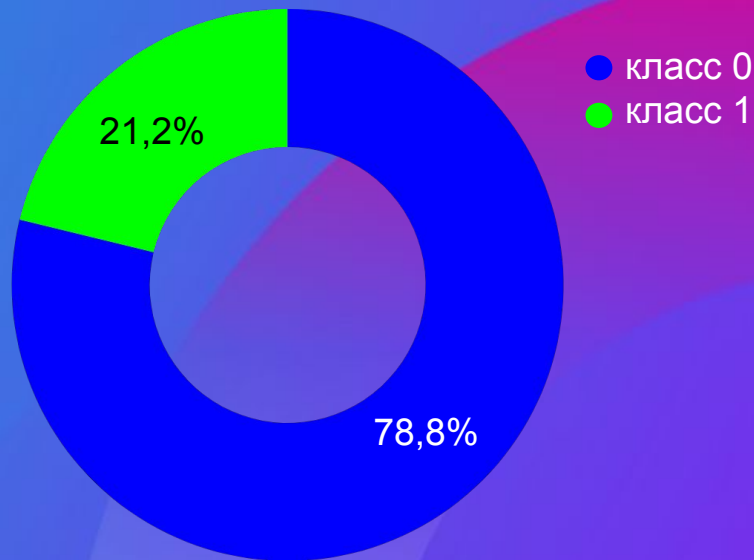
# АНАЛИЗ ИСХОДНЫХ ДАННЫХ

- 28 901 примеров
- 22 788 хороших (0)
- 6 113 плохих (1)

Дисбаланс в сторону первого класса

Данные – списки json с пятью последними фразами и таргет фраза, следующая после контекста

```
[{'user': 'маруся через десять минут будильник включи',  
  'marusia': 'Окей, разбужу вас через 10 минут'},  
{ 'user': 'тормози', 'marusia': ""},  
{ 'user': 'будильник на три часа ночи',  
  'marusia': 'Окей, разбужу вас через 1 час 30 минут'},  
{ 'user': "",  
  'marusia': ' [Доброй ночи, Алексей!] | [date: Воскресенье, 9 октября] | [temp: 12] |  
[ Погода {Погода в выходные}, Давай сыграем в города {Давай сыграем в  
города}, Радио {Поставь Радио Энерджи}, Таймер {Таймер на 30 секунд}}'],  
{ 'user': 'привет', 'marusia': 'Привет, Алексей!'}]
```



# АНАЛИЗ И ПОДГОТОВКА ДАННЫХ

Очистка диалогов от  
эмодзи

"🙄 Это Мистер Декарт."  
"👏 Благодарю за игру!"  
"😬 Странно, что я ошиблась"



"Это Мистер Декарт."  
"Благодарю за игру!"  
"Странно, что я ошиблась"

Восстановление  
пунктуации

"маруся я очень тебя люблю"  
"марусь включи пожалуйста  
елена вальяк гипноз для сна"



"Маруся, я очень тебя люблю."  
"Марусь включи, пожалуйста,  
Елена Вальяк. Гипноз для сна."

Поиск фраз с  
латиницей

"выключи bluetooth"  
"включи led zeppelin лестница в небеса"

Пустые фразы в  
контекстах

Только в 4% контекстов больше половины фраз  
отсутствуют

# ПРИЗНАКИ ДЛЯ МОДЕЛИ

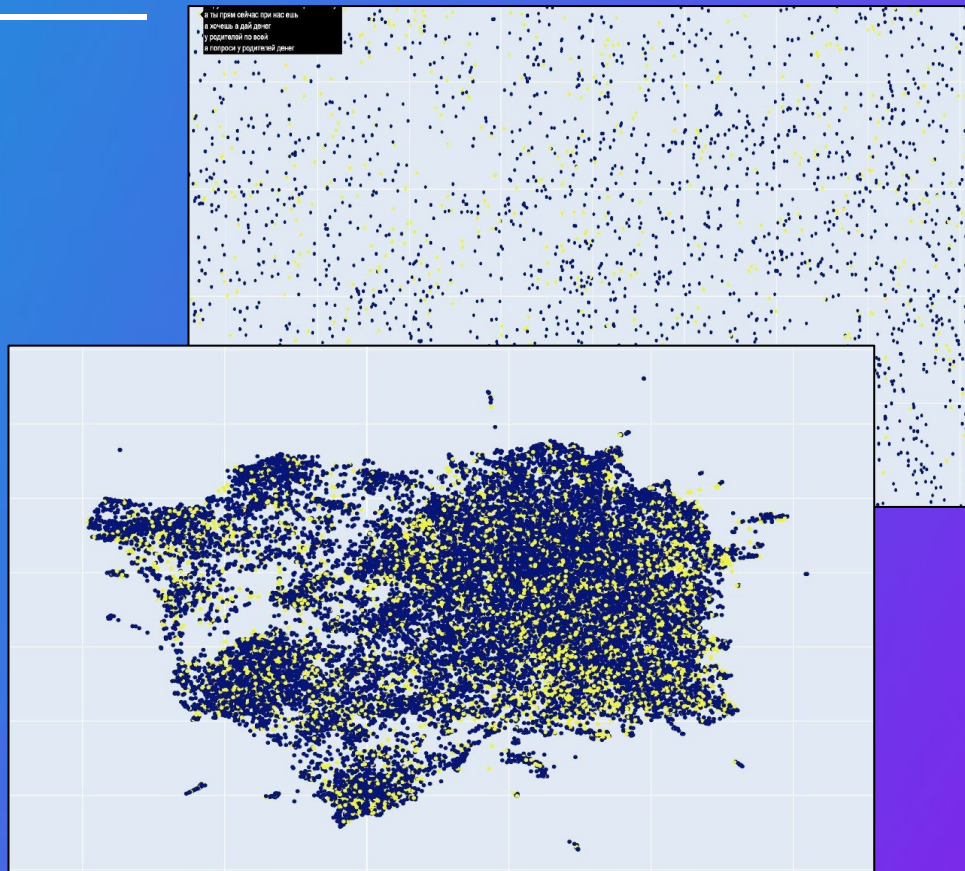
Признаки нужны для улучшения качества модели. Оценивали влияние признаков их корреляцией с лейблом – то есть насколько и в какую сторону изменение признака влияет на лейбл

ПРИЗНАК	КОРРЕЛЯЦИЯ
Длина таргет фразы	-0.16
Метрика Relevance на диалоге пользователя с Марусей	-0.15
Метрика Readability на таргет фразах	-0.14
Наличие именованных сущностей (имен, названий, адресов и т.д.)	-0.11
Метрика Linguistic acceptability (приемлемость высказывания)	-0.06
Число пустых / не пустых фраз в контекстах	-0.05
Метрика токсичности по фразам и контекстам пользователя	-0.02
Количество подряд идущих пустых фраз в контекстах	0.00
Наличие шаблонных фраз Маруси (про погоду и т.д.)	0.00
Пропуск Марусей последней фразы в диалоге при наличии фразы пользователя (при просьбах сделать потише / выключить музыку)	0.00

# ГИПОТЕЗА: ТЕМАТИКИ ДИАЛОГОВ

1. Для каждого примера получили его смысловой эмбединг модели MUSE (набор чисел, приблизительно отражающий смысл текста)
2. Перевели эмбединги текстов в двумерное пространство (умар) и нарисовали их
3. Окрасили примеры соответственно лейблам

Заметно, что зависимости между лейблом и тематикой нет, значит фича плохая



# АУГМЕНТАЦИЯ ПРИМЕРОВ

Из-за дисбаланса в сторону 0 класса (примеров 1 класса всего 20%) решили попробовать сгенерировать дополнительные примеры и добавить в обучающую выборку

Перемешать плохие контексты с плохими фразами

1. Берем контексты плохого класса с низкой метрикой релевантности
2. Берем таргет фразы плохого класса
3. Мешаем их между собой и получаем новые примеры

RuDialogPT-2 на плохих контекстах

1. Берем модель для генерации текстов
2. Дообучаем ее на контекстах
3. Генерируем контексты и подаем в обучающую выборку добавив случайные плохие таргеты

# АУГМЕНТАЦИЯ ПРИМЕРОВ ГЕНЕРАТИВНОЙ МОДЕЛЬЮ

```
[{'user': 'Какая погода в Москве?',  
  'marusia': '<Показываю анимацию в зоне часов> сегодня в Москве будет облачно, дождь температура от шести'},  
 {'user': 'На машине.', 'marusia': 'На машине времени?'},  
 {'user': 'Да.', 'marusia': 'Вот это да'},  
 {'user': 'Ты меня любишь?', 'marusia': 'Это вы мне?'},  
 {'user': 'Ты любишь меня?', 'marusia': 'Я вас очень люблю!'}]
```

```
[{'user': 'Маруся, поставь музыку',  
  'marusia': 'vk: Авторизуйтесь через VK ID и слушайте миллионы треков'},  
 {'user': 'Маруся, включи.', 'marusia': 'Ставлю Король и Шут.'},  
 {'user': 'Давай.', 'marusia': 'Ставлю Король и Шут.'},  
 {'user': 'Давай, пока.', 'marusia': ''},  
 {'user': '', 'marusia': ''}]
```

```
[{'user': 'Маруся, как дела?', 'marusia': 'Да вроде неплохо'},  
 {'user': 'Маруся, не спишь?', 'marusia': 'Да вроде нет, у меня дел полно'},  
 {'user': 'Маруся, давай поиграем в "Кто хочет стать миллионером".',  
  'marusia': 'А давайте'},  
 {'user': 'Да.', 'marusia': 'Ну супер'},  
 {'user': 'Маруся, давай поиграем в «Кто хочет стать миллионером».',  
  'marusia': 'Давайте сыграем! Я хочу стать миллионером. На этот раз я хочу стать миллионером!'}]
```

Дообучали <https://huggingface.co/tinkoff-ai/ruDialoGPT-medium>  
Но использовать не успели :(



# МОДЕЛЬ И ЕЕ ОБОСНОВАНИЕ

Использовали для дообучения трансформерные модели, зарекомендовавшие себя в похожих задачах. Замеряли на тестовой подвыборке из тренировочного датасета

**sberbank-ai/ruBert-large**

**0.75**

**DeepPavlov/xlm-roberta-large-en-ru**

**0.78**

**sberbank-ai/ruT5-large**

**0.70**

**DeepPavlov/xlm-roberta-large-en-ru**

# ЧТО ПРОБОВАЛИ

## Делили на три выборки:

- train (70%)
- validation (20%)
- test (10%)

Пробовали делить на фолды, но качество не было удовлетворительным

## Модели:

- sberbank-ai/ruBert-large
- sberbank-ai/ruT5-large
- DeepPavlov/xlm-roberta-large-en-ru
- xlm-roberta-large

Остановились на небольших, потому что другие обучались медленнее и показывали результат хуже

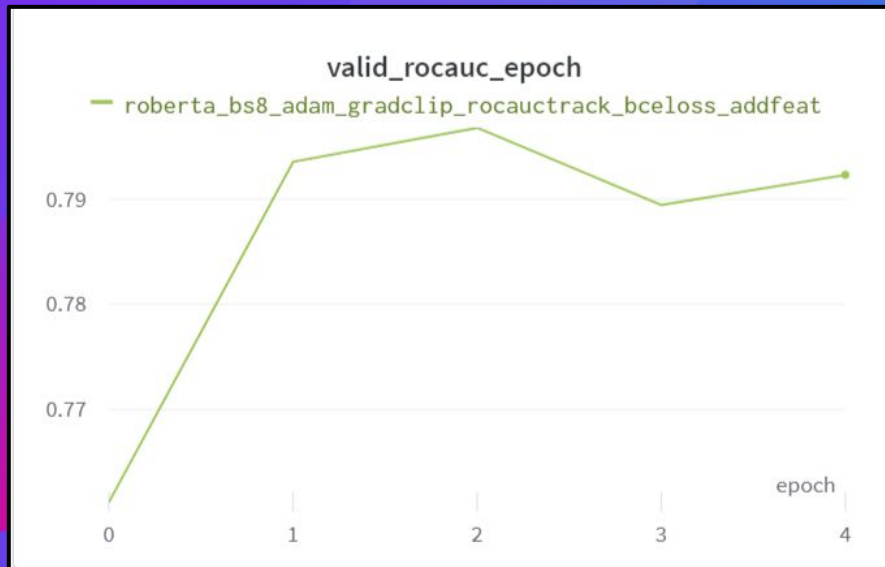
## Спецтокен:

Добавили в токенизатор специальный токен [RESPONSE\_TOKEN] для разделения контекста от таргет-фразы и не получили профита

## Сэмплинг

Искусственные примеры подняли метрику

# РЕЗУЛЬТАТЫ МОДЕЛИ



ROC-AUC: 0.79



LOSS: 0.39

# ВАРИАНТЫ РАЗВИТИЯ РЕШЕНИЯ И РЕАЛИЗАЦИЯ

Готовую модель можно заhostить через streamlit с интерфейсом от huggingface spaces

Маруся не отвечает на реплики

из телевизора

## VK кейс Маруся GrowAi



📄 - Об этом приложении

## Краткое описание решения

Мы представляем алгоритм машинного обучения для решения задачи бинарной классификации, отличающий команды пользователя голосового помощника VK "Маруся" от внешнего шума для повышения коммуникации.

Проанализировав входные данные нами было проведена комплексная оценка параметров признаков, в процессе подготовки было очищено порядка 1,5 тысяч диалогов, содержащих

# маруся

📄 - Об этом приложении

## Краткое описание решения

Мы представляем алгоритм машинного обучения для решения задачи бинарной классификации, отличающий команды пользователя голосового помощника VK "Маруся" от внешнего шума для повышения коммуникации.

Проанализировав входные данные нами было проведена комплексная оценка параметров признаков, в процессе подготовки было очищено порядка 1,5 тысяч диалогов, содержащих смайлики и пустые значения для лучшего распознавания команд алгоритмом, проставлена пунктуация в ключевых фразах и диалогах пользователей, дана оценка токсичности пользовательских ответов, качества данного ответа, а также лингвистическая приемлемость.

Стек решения: python, torch, transformers, sklearn, scipy, pandas, numpy, streamlit

Уникальность: Решение оформлено в виде web приложения с пользовательским интерфейсом, позволяющим загружать релевантные данные для их обработки разработанным алгоритмом. Интерпретируемость модели определяется извлечением отдельных признаков, имеющих наибольшую взаимосвязь и значимость для классификации. Повышение точности основано на подходе разделения дата сета на 5 частей для обучения отдельных моделей вместо 1.

позволяющим загружать релевантные данные для их обработки разработанным алгоритмом. Интерпретируемость модели определяется извлечением отдельных признаков, имеющих наибольшую взаимосвязь и значимость для классификации. Повышение точности основано на подходе разделения дата сета на 5 частей для обучения отдельных моделей вместо 1.

## Проверить работу алгоритма, можно через эту форму

Вставьте или напишите текст

🌟 ТыК



Drag and drop file here  
Limit 200MB per file

Browse files



Или вы можете загрузить сюда файл

# ТЕХНОЛОГИИ



# Transformers

---

# КОМАНДА

## GrowAi

---



**Елисей Рыков**

Tg: @Imeribal



**Александр Туманов**

Tg: @ftfy\_m



**Анна Белькова**

Tg: @GRECHAX