

# DATA FEST 2023

20.05.2023 – 04.06.2023

Тема доклада:

Создание синтетического датасета для SFT  
LM на задачи суммаризации вакансий

Спикер:

**Никита Венедиктов**



# Познакомимся?



Магистрант AI Talent Hub ITMO

Сейчас НЛПишник в биоинформатическом стартапе  
Deeploid

Раньше разрабатывал системы для телекома  
в компании Keysight Technologies

В команде работа.ру участвую в улучшении сервиса  
по суммаризации вакансий



**I/ITMO**

**AI TALENT HUB**

**работа.ру**

**Никита  
Венедиктов**

Команда улучшает сервис суммаризации вакансий → повышается конверсия

Решили делать это на основе дообучения разных LLM

**Data Scientist**  
по договоренности

В команду «Гео и Графы» кластера «Data Science и клиентская аналитика» корпоративного блока Сбербанка мы ищем специалистов по работе с данными. Вам предстоит работать с графами связей, строить витрины данных, использующиеся

СБЕР  
Россия, Санкт-Петербург

Откликнуться

СБЕР БАНК

**Data scientist/ML-инженер**  
по договоренности

Rambler&Co занимает первое место среди медиахолдингов России по объему аудитории цифровых ресурсов — свыше 45,2 млн человек ежемесячно. В него входит «Лента.ру», «Газета.Ru», Championat.com, портал «Рамблер», «Секрет фирмы» и ряд

Rambler&Co

Откликнуться

RAMBLER&Co

суммаризация

саммаризация

парафраз



# Процесс обучения

Phase 1. Pretraining for completion

Phase 2. **Supervised finetuning (SFT)**  
to summarize vacancies

Phase 3. RLHF

3.1. Reward model (RM)

3.2. Finetuning using the reward model

Step 1  
**Collect demonstration data,  
and train a supervised policy.**

A prompt is  
sampled from our  
prompt dataset.

🧠  
Explain the moon  
landing to a 6 year old

A labeler  
demonstrates the  
desired output  
behavior.

👤  
Some people went  
to the moon...

This data is used  
to fine-tune GPT-3  
with supervised  
learning.

SFT  
🧠  
📄📄📄

Step 2  
**Collect comparison data,  
and train a reward model.**

A prompt and  
several model  
outputs are  
sampled.

🧠  
Explain the moon  
landing to a 6 year old

A B  
Explain gravity... Explain war...  
C D  
Moon is natural satellite of People went to the moon...

A labeler ranks  
the outputs from  
best to worst.

👤  
D > C > A = B

This data is used  
to train our  
reward model.

RM  
🧠  
D > C > A = B

Step 3  
**Optimize a policy against  
the reward model using  
reinforcement learning.**

A new prompt  
is sampled from  
the dataset.

🐸  
Write a story  
about frogs

The policy  
generates  
an output.

PPQ  
🧠

Once upon a time...

The reward model  
calculates a  
reward for  
the output.

RM  
🧠

The reward is  
used to update  
the policy  
using PPO.

$r_k$

# Хватит ли RLHF? Не полетело

Step 2

Collect comparison data,  
and train a reward model.

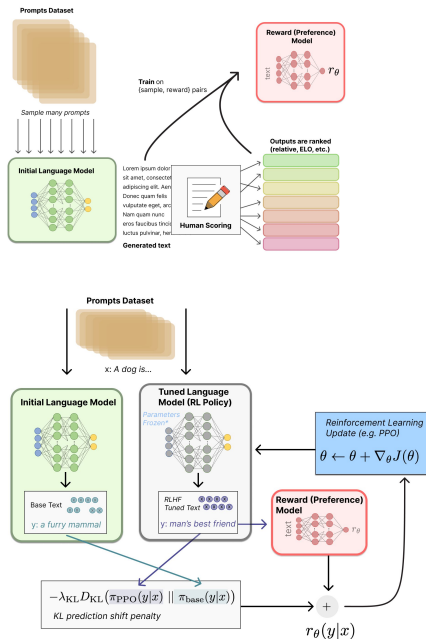
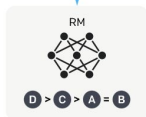
A prompt and  
several model  
outputs are  
sampled.



A labeler ranks  
the outputs from  
best to worst.



This data is used  
to train our  
reward model.



Своими силами, немного дообучим и дадим  
уже пользователям пользователям  
возможность ранжировать выдачи модели

-неа



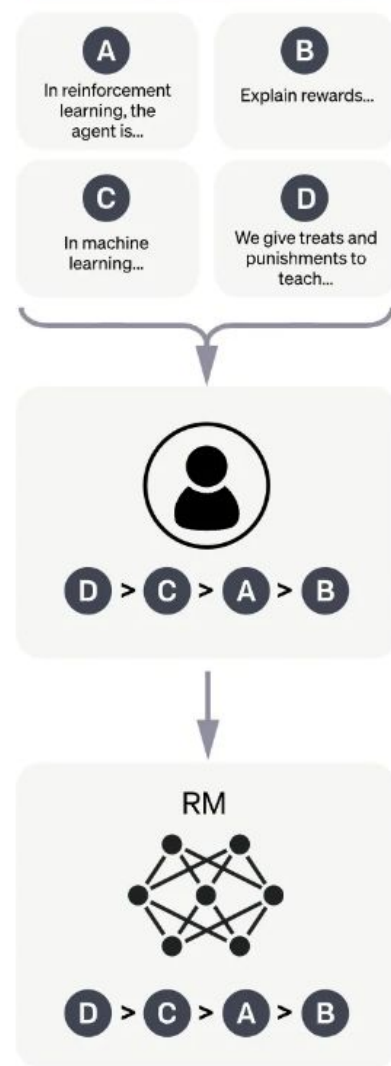
- Особенности предобучения (новости, статьи)
- HR домен
- Галюцинации

# Итоговое решение

1. Выбор трансформера (mBART, T5, GPT)
2. Синтезируем датасет для адаптации под наши задачи
3. Этап **Supervised finetuning (SFT)**
4. RLHF - используя модель с 3го этапа генерируем и первый раз руками размечаем датасет для reward модели 2-3к саммарии
5. В продакшн и далее используем отклики пользователей для формирования большого датасета обучения reward модели
6. Подводим итоги

Важно что мы не сразу в автоматический продакшн

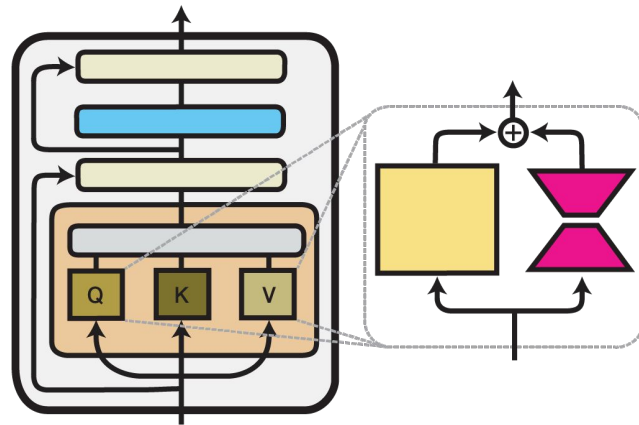
А сначала предложим работодателям несколько вариантов, чтобы выбрали сами и помогли нам



# Синтезировать датасет для SFT

Supervised finetuning (SFT) позволяет сначала адаптировать модель под наши задачи (вешаем на LLM Lora-адаптер для нашего домена ) но для этого нужен датасет

Датасет вакансий есть, осталось его изучить и найти оптимальный метод суммаризации





# Особенности вакансий — проблемы?

- проблема в экстракции сути — истории компаний
- невозможность ранжирования предложений в вакансии — однострочник
- проблема в точности экстракции — если прям ладненько, но много
- надо было адаптировать предобработку — много цифр
- нарезать маркированные списки на предложения

Слова, словосочетания, предложения?

Потом попросим LLM их перефразировать и будет меньше галлюцинаций

Чем их вытягивать из текста?



# Суммаризация — база

## Input Article

Marseille, France (CNN) The French prosecutor leading an investigation into the crash of Germanwings Flight 9525 insisted Wednesday that he was not aware of any video footage from on board the plane. Marseille prosecutor Brice Robin told CNN that "so far no videos were used in the crash investigation." He added, "A person who has such a video needs to immediately give it to the investigators." Robin's comments follow claims by two magazines, German daily Bild and French Paris Match, of a cell phone video showing the harrowing final seconds from on board Germanwings Flight 9525 as it crashed into the French Alps. All 150 on board were killed. Paris Match and Bild reported that the video was recovered from a phone at the wreckage site. ...

## Text Summarization Models

Abstractive summarization

Extractive summarization

## Generated summary

Prosecutor : " So far no videos were used in the crash investigation "

## Extractive summary

marseille prosecutor brice robin told cnn that " so far no videos were used in the crash investigation . " robin 's comments follow claims by two magazines , german daily bild and french paris match , of a cell phone video showing the harrowing final seconds from on board germanwings flight 9525 as it crashed into the french alps . paris match and bild reported that the video was recovered from a phone at the wreckage site .

Автоматическое создание краткого содержания

**Экстрактивная** — это когда для суммаризации используются кусочки оригинального текста в том виде, как они были написаны изначально.

**Абстрактивная** — это генеративная модель, которая создает абсолютно новый текст, где могут содержаться слова, которые не встречались в оригинальном тексте.

В сервисах ты настраиваешь, количество слов, токенов, предложений

Или соотношение вход к выходу



# Абстрактивная

Нейронные подходы к абстрактной суммаризации концептуализируют задачу как (seq2seq или text2text)

И сразу приходят в голову решение — трансформеры

А именно функция саммарайз. Но вот только есть свои приколы)

Тяжелые, галюцинации и “неправильный” эттеншн из-за особенностей предобучения



# Нас таким не запугать

Тяжесть этим — бьем этими приемами (Квантизация, Прунинг, Дистилляция)

Кривые результаты из-за особенностей обучения — плохо с вакансиями работает дообучения — бьем SFT + RLHF — но его пока нет)

Галлюцинации — они критичны на этом этапе

Это из пушки по воробьям и тяжело контролировать  
Пока я это отложил, но использовал потом для перефразирования

# Экстрактивная суммаризации предложений

Методы данного подхода характеризует наличие оценочной функции важности информационного блока. Ранжируя эти блоки по степени важности и выбирая ранее заданное их число, мы формируем итоговое резюме текста.

- Graph based (TextRank, LexRank- на основе pagerank)
- Fuzzy logic based
- Topic Words (Luhn)
- Latent Semantic Analysis Method (LSA)
- Frequency-driven (Tf-IDF, KL Divergence)

Table 3 – Average context score for the keyword extraction models based on the original text

Method	Average context score
Textrank - BERT	<b>0,5030</b>
K-means - BERT	0,4727
K-means - TF-IDF	0,4555
Textrank - TF-IDF	0,4546
LSA	0,3893
Random baseline	0,3612



# Про библиотеки

Тестил низкоуровневые реализации - требуют допиливания, хотя много и есть лекции

Нашел либу — совладал

```
from sumy.nlp.tokenizers import Tokenizer
from sumy.summarizers.luhn import LuhnSummarizer
from sumy.summarizers.text_rank import TextRankSummarizer
from sumy.summarizers.lsa import LsaSummarizer
from sumy.summarizers.lex_rank import LexRankSummarizer
from sumy.summarizers.kl import KLSummarizer
from sumy.nlp.stemmers import Stemmer
```



# Результаты слепого тестирования

Команда голосовала и вышли в лидеры:

- Пайплайн с `sumy_TextRank`
- Пайплайн с `sumy_LSA`

Немного про них



# Textrank

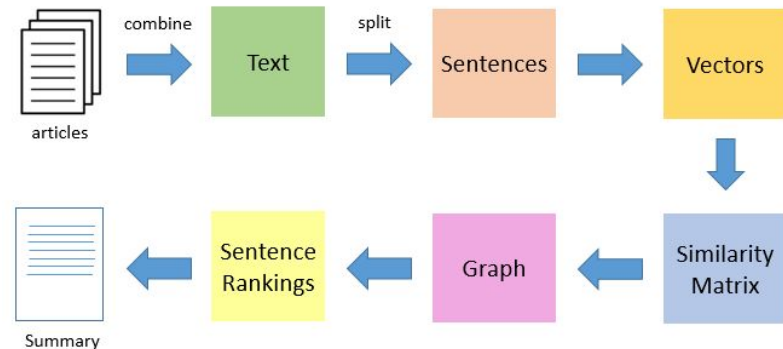
Оценить сходство между предложениями используя матрицу сходства, рассчитанную с помощью косинусного сходства. Матрица сходства затем используется в качестве весов для ребер графа.

Важным аспектом TextRank является то, что он не требует глубоких лингвистических знаний, аннотированных корпораций, специфичных для конкретной области или языка, что делает его легко переносимым на другие области, жанры или языки.

1. Первым шагом будет конкатенация всего текста, содержащегося в вакансии.
2. Затем разделите текст на отдельные предложения.
3. На следующем этапе мы находим векторное представление (вставки слов) для каждого предложения.
4. Сходство между векторами предложений вычисляется и сохраняется в матрице.
5. Позже матрица сходства преобразуется в граф, где предложения являются вершинами, а оценки сходства - ребрами, для вычисления диапазона предложений.
6. Наконец, определенное количество предложений с наилучшим рейтингом формирует окончательное резюме.

## Weighted Graphs

$$WS(V_i) = (1 - d) + d * \sum_{V_j \in In(V_i)} \frac{w_{ji}}{\sum_{V_k \in Out(V_j)} w_{jk}} WS(V_j)$$



of two sentences with the length of each sentence. Formally, given two sentences  $S_i$  and  $S_j$ , with a sentence being represented by the set of  $N_i$  words that appear in the sentence:  $S_i = w_1^i, w_2^i, \dots, w_{N_i}^i$ , the similarity of  $S_i$  and  $S_j$  is defined as:

$$Similarity(S_i, S_j) = \frac{|\{w_k | w_k \in S_i \& w_k \in S_j\}|}{\log(|S_i|) + \log(|S_j|)}$$





# Latent Semantic Analysis (LSA)

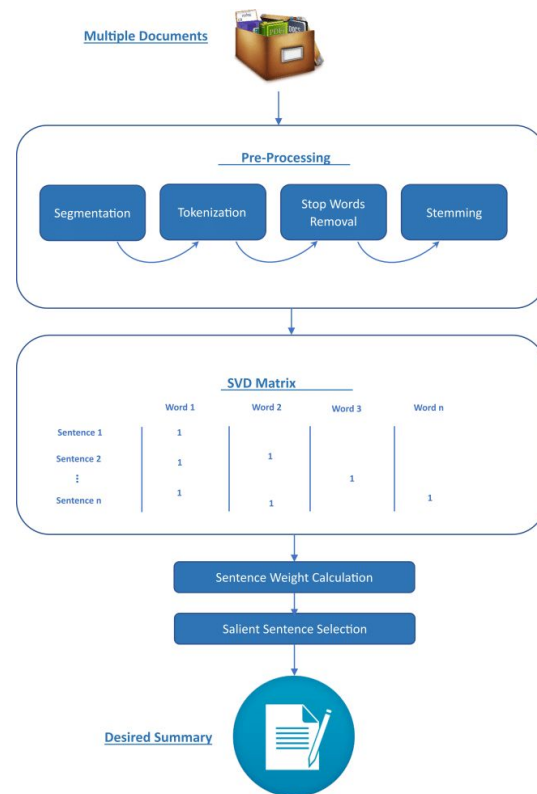
LSA опирается на статистические методы для оценки скрытой семантической структуры данных. LSA строит латентное семантическое пространство, в котором документы и слова представлены в виде векторов.

LSA использует математическую технику Singular Value Decomposition (SVD) для распознавания связей между различными терминами и понятиям.

Гонг и Х. Лю (2001) предложили метод обобщения, который использует LSA для того для создания резюме на основе наиболее важных предложений.

# Latent Semantic Analysis (LSA)

1. Предварительная обработка текста:
2. Построение матрицы термов-документов:  
Строится матрица, где строки представляют термины (слова), а столбцы – документы.  
В ячейках матрицы указывается частота встречаемости термов в каждом предложении или, возможно, более сложные меры веса, такие как TF-IDF.
3. Применение сингулярного разложения (SVD):
  - Матрица  $U$  представляет собой матрицу левых сингулярных векторов и содержит информацию о семантической структуре термов.
  - Матрица  $S$  – диагональная матрица, содержащая сингулярные значения, отражающие важность скрытых семантических факторов.
  - Матрица  $V$  представляет собой матрицу правых сингулярных векторов и содержит информацию о семантической структуре документов.
4. Сокращение размерности: – Это позволяет удалить шум и сохранить наиболее информативные и семантически значимые факторы.
5. Вычисление семантической близости:  
Используя усеченные матрицы  $U$  и  $S$ , можно вычислить близость между терминами или документами





# Yes! — результаты экстрактивный

пока 20к синтезированных для SFT

Вакансия

Требования: Строительная компания приглашает на работу вахтовым методом. Вахта на Север 60/30. Заработная плата 150 тыс. рублей. Обратите внимание: опыт работы в аналогичной должности строго ОБЯЗАТЕЛЕН. Оформление по ТК РФ, з/п 2 раза в месяц на карту. Работодатель предоставляет проживание, 3-разовое питание, спецодежду, обувь. Компенсация проезда  
Информация о работодателе: Строительная компания.

sumy\_LexRank  
Строительная компания приглашает на работу вахтовым методом. Заработная плата 150 тыс. рублей. Обратите внимание.

sumy\_Luhn Строительная компания приглашает на работу вахтовым методом. опыт работы в аналогичной должности строго ОБЯЗАТЕЛЕН. Строительная компания.

sumy\_KLdiv  
Строительная компания приглашает на работу вахтовым методом. Вахта на Север 60/30. з/п 2 раза в месяц на карту.

sumy\_LSA  
Строительная компания приглашает на работу вахтовым методом. опыт работы в аналогичной должности строго ОБЯЗАТЕЛЕН. з/п 2 раза в месяц на карту.

sumy\_TextRank  
Строительная компания приглашает на работу вахтовым методом. з/п 2 раза в месяц на карту. Работодатель предоставляет проживание.



# Про модель перефразирования

Выбрали cointegrated/rut5-base-paraphraser

```
from transformers import T5ForConditionalGeneration, T5Tokenizer  
MODEL_NAME = 'cointegrated/rut5-base-paraphraser'  
model = T5ForConditionalGeneration.from_pretrained(MODEL_NAME)  
tokenizer = T5Tokenizer.from_pretrained(MODEL_NAME)
```

Выбрали экстрактивные реализации от 200-300 символов  
и перефразируем их в 150-200 символов

Настроил модель перефразирования

протестил - результат порадовал, но уже сильно дольше



# Yes! — результаты перефразированных

## Результаты

sumy\_LSA

Строительная компания приглашает на работу вахтовым методом. опыт работы в аналогичной должности строго ОБЯЗАТЕЛЕН. з/п 2 раза в месяц на карту.

para\_LSA

Компания "Строительная компания" приглашает вахтовиков на постоянную работу. З/п не менее 2 раз в месяц.

sumy\_TextRank

Строительная компания приглашает на работу вахтовым методом. з/п 2 раза в месяц на карту. Работодатель предоставляет проживание.

para\_TextRank

Компания строителей приглашает на вахту. З/п от 2 раз в месяц. Работодателю предоставляется жилье.





# Yes! — результаты перефразированных

20к + 20к\* - 40к синтезированных для SFT

\* мы уже плюсуем абстрактный вариант

После дообучения LLM не должна хакнуть алгоритмы экстрактивной суммаризации, потому что мы использовали парафрайзер

убили 2 зайцев: сжимать и работать абстрактно



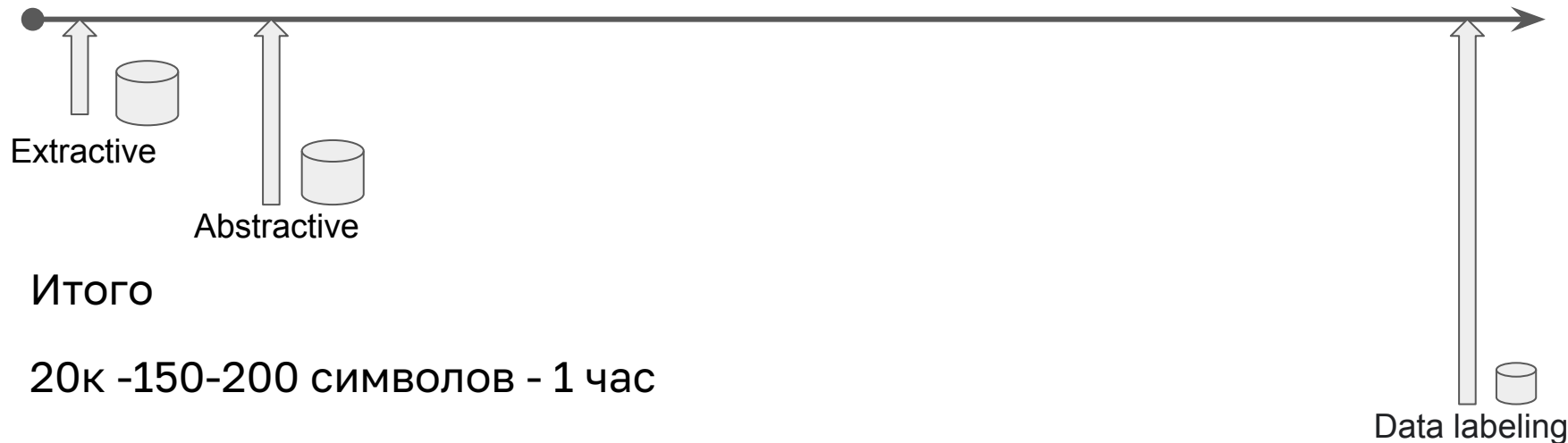
# Yes! — результаты всех

20к +20к+7к\* - 47к синтезированных для SFT

Со звездочкой это уже: ручками

Ничего интересного, только тяжелый труд

# ДАТАСЕТ ДЛЯ SFT



Итого

20к -150-200 символов - 1 час

20к -150-200 символов (до 53 токенов) - 1 день

7к - 150-200 символов (краусорс) - 1 месяц





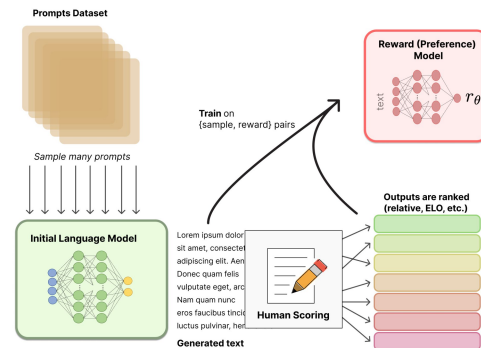
# Итоги

- Датасет составлен на основе HR домена
- Содержит большинство важной инфы с точки зрения работодатель-соискатель
- Содержит экстрактивную, абстрактивную и ручную разметку

Позволяет дообучить через SFT любой трансформер

# Что было дальше? Продакшн?

1. Модель после SFT генерирует датасет для RLHF
2. Этап RLHF
  - показываем несколько саммари и даем окошко для своего
  - вручную экспертами ранжируем 1-1.5 к саммари
  - Сохраняем данные о выборе
  - обучаем RM
3. Так устроили бизнес процессы, что можем использовать RLHF модель итеративно





# Спасибо!



@SPACE\_APPLE

tg, inst: @space\_apple

**ИТМО**

**AI TALENT HUB**

**работа.ру**



**Никита  
Венедиктов**



# Краткое описание задачи RLHF и ее значимость

## Reinforcement Learning from Human Feedback

*обучение с подкреплением на основе человеческих предпочтений*

ну это база - обратная реакция от пользователя

ВОТ В GPT

The screenshot shows a dark-themed dialog box titled "Provide additional feedback" with a close button (X) in the top right corner. Inside the dialog, there is a text input field with the placeholder text "What was the issue with the response? How could it be improved?". Below the input field, there are three checkboxes with corresponding labels: "This is harmful / unsafe", "This isn't true", and "This isn't helpful". At the bottom right of the dialog, there is a button labeled "Submit feedback".