

# Self-Instruct

Черномордин Родион 211

# В нашем мире (весна 2023):

14 марта выстреливает ChatGPT

Проблемы:

- дорого и тяжело сделать инструкции на все задачи в мире
- требует творческого подхода
- опыт для решений к каждой задачи

**Self-Instruct** - метод, помогающий LM улучшать их способность исполнять инструкции на естественном языке, где модель обучается на основе



# Self-Generated Instructions

Статья 1

# Инструкции:

- Набор инструкций с задачами  $t \in \{I_t\}$
- Задача  $t$  имеет  $n_t \geq 1$  экземпляров  $\{(X_{t,i}, Y_{t,i})\}_{i=1}^{n_t}$ , где  $M(I_t, X_{t,i}) = Y_{t,i}$

## Пример:

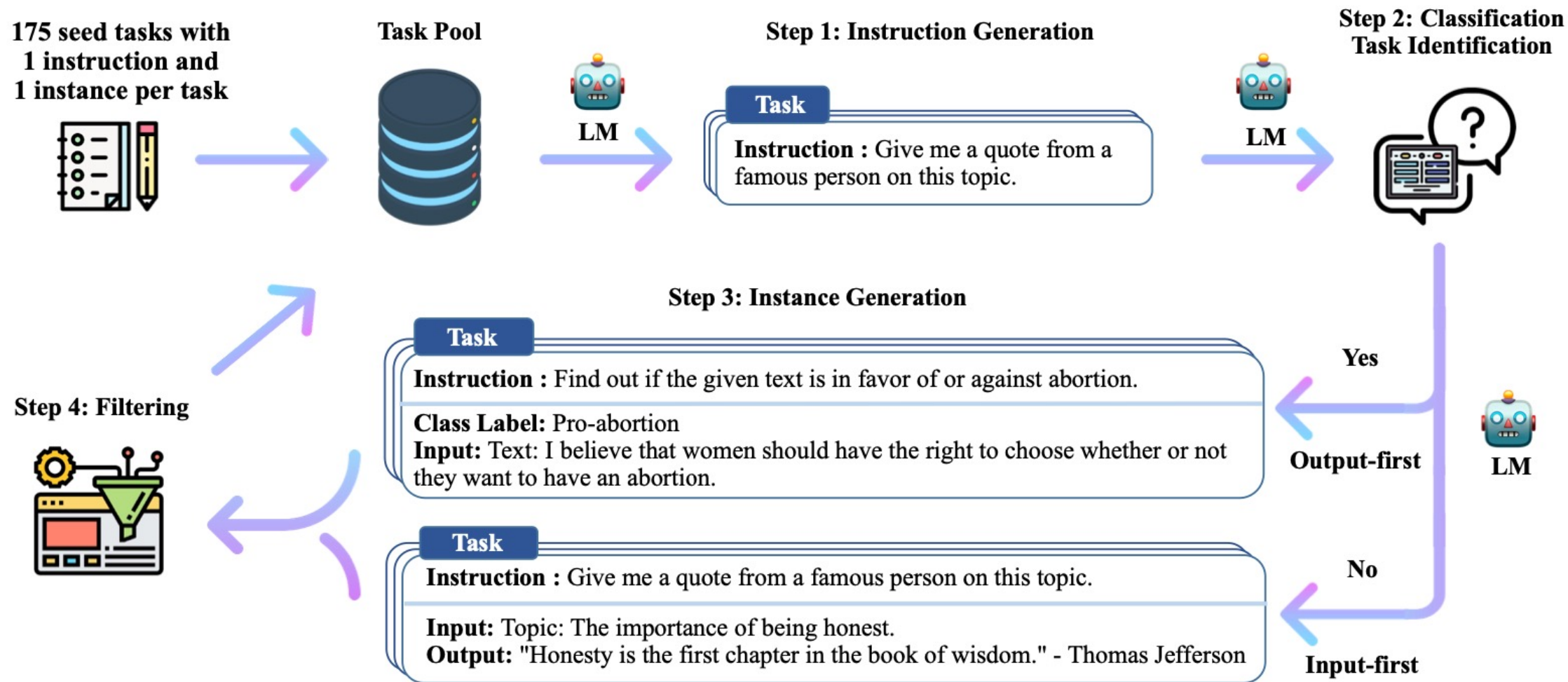
$I =$  “подготовь доклад о Self-Instruct ”

$X =$  “”

$I =$  “подготовь доклад на следующую тему ”

$X =$  “Self-Instruct ”

# Конвейер генерации данных



# Шаг 1 - генерация

175 seed tasks with  
1 instruction and  
1 instance per task



Task Pool



LM

Step 1: Instruction Generation

Task

**Instruction :** Give me a quote from a famous person on this topic.

Всего берется 8 задач:

- 6 – написанные человеком
- 2 – сгенерированные на предыдущих итерациях

Come up with a series of tasks:

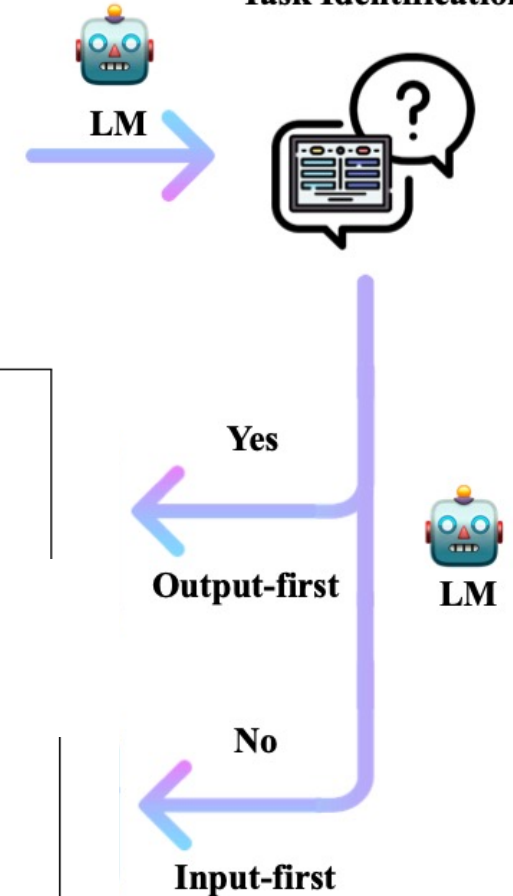
```
Task 1: {instruction for existing task 1}
Task 2: {instruction for existing task 2}
Task 3: {instruction for existing task 3}
Task 4: {instruction for existing task 4}
Task 5: {instruction for existing task 5}
Task 6: {instruction for existing task 6}
Task 7: {instruction for existing task 7}
Task 8: {instruction for existing task 8}
Task 9:
```

# Шаг 2: классификация данных

31 задача из стартового набора данных:

- 12 не из классификации
- 19 из классификации

## Step 2: Classification Task Identification



Can the following task be regarded as a classification task with finite output labels?

Task: Given my personality and the job, tell me if I would be suitable.  
Is it classification? Yes

...

Task: Given a set of numbers, find all possible subsets that sum to a given number.  
Is it classification? No

Task: {instruction for the target task}



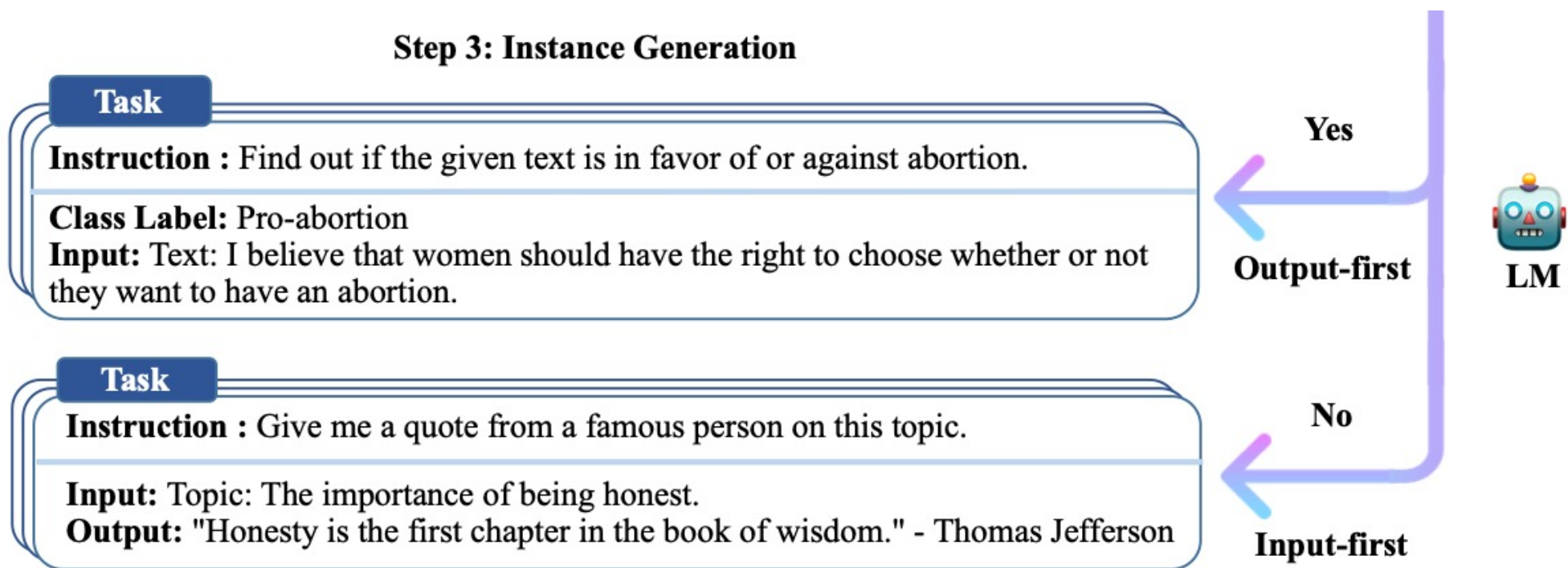
# Шаг 3: создание примеров

Input-first и классификация:

**Instruction:** найди в каком слове есть грамматические ошибки

**Input:** ча ща жи ши

**Output:** 3



# Шаг 4: чистка

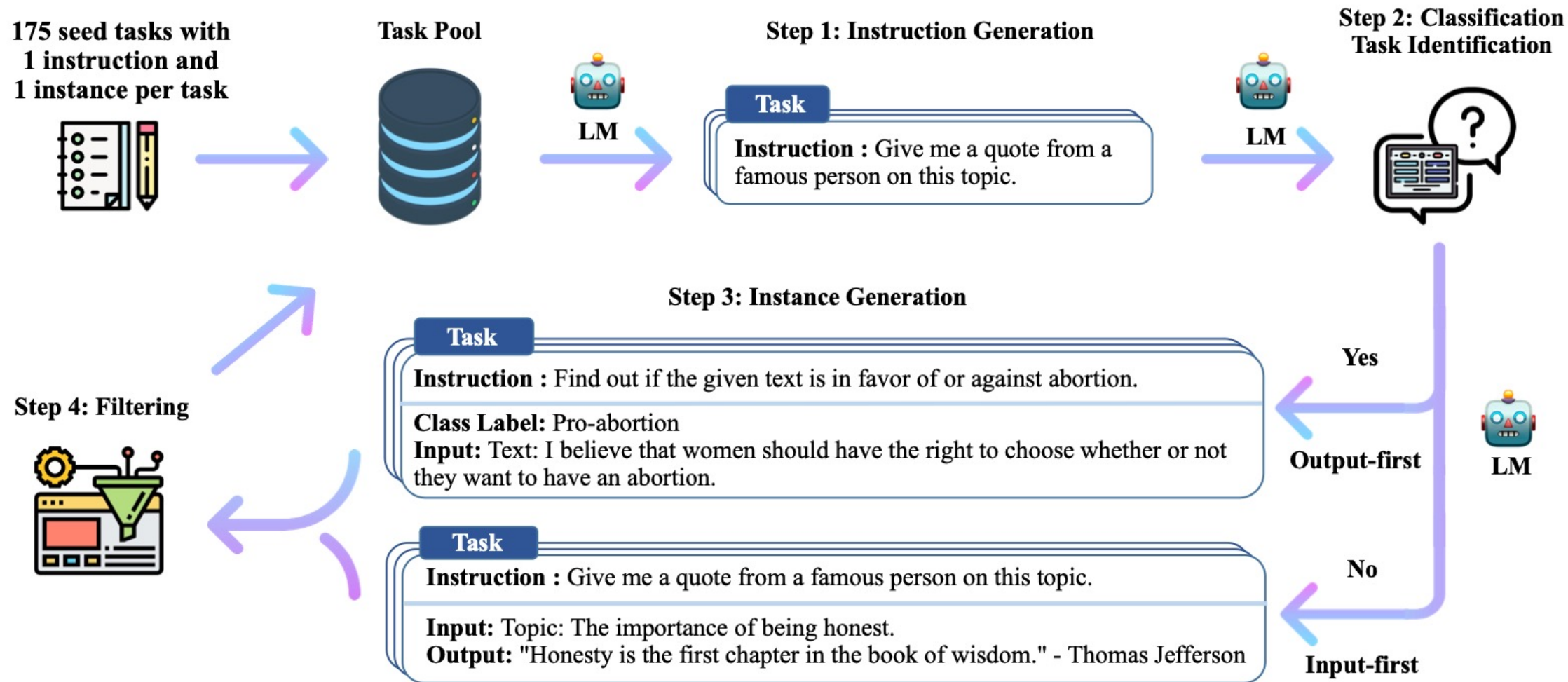
## Step 4: Filtering



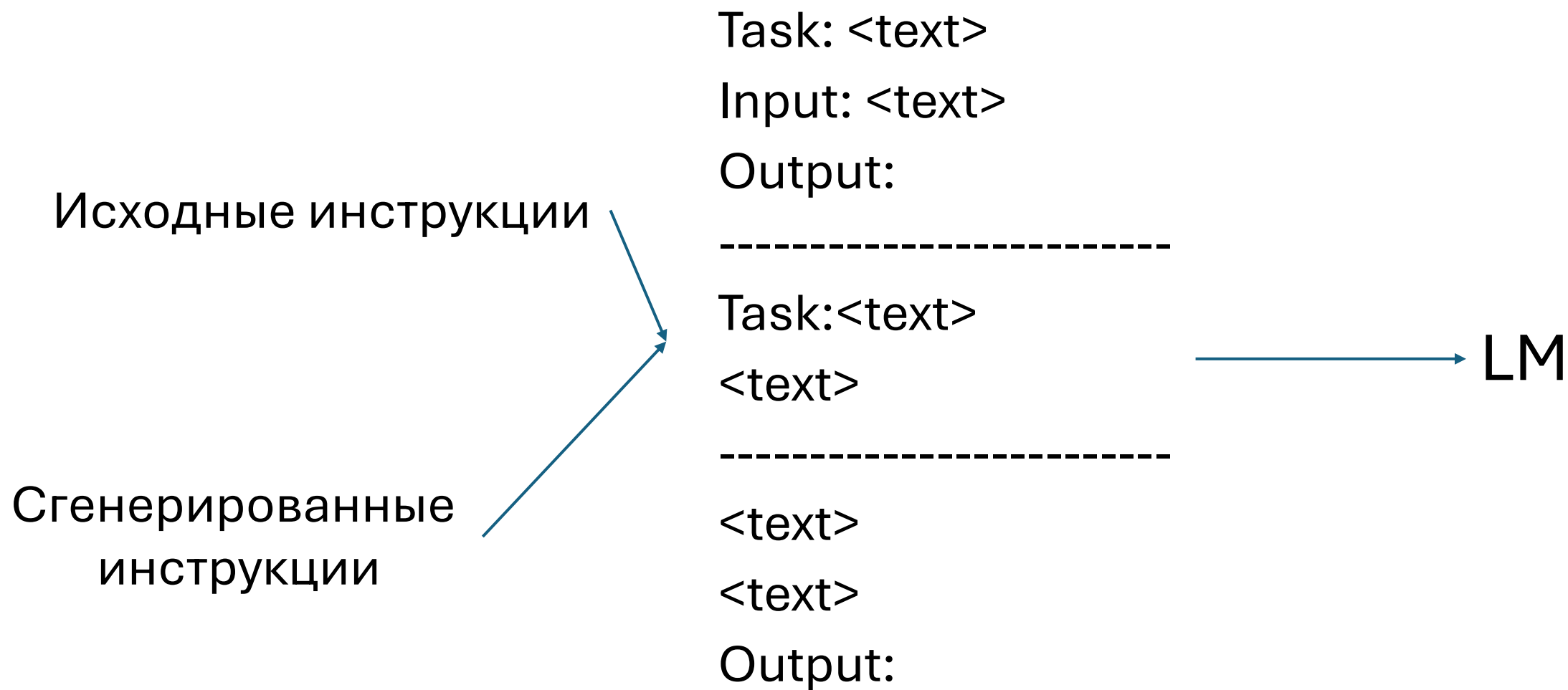
Удаляем данные, которые:

1. Имеют ROUGE-L (Recall-Oriented Understudy for Gisting Evaluation by Longest Common Subsequence)  $> 0.7$
2. Инструкции про изображения, графы и тд
3. Точно такие же инструкции или входные данные, но разные выходы
4. Эвристики: слишком длинные/короткие, выходные данные повторения входных и тд

# Конвейер генерации данных



# Self-Instruct



# GPT3

statistic	
# of instructions	52,445
- # of classification instructions	11,584
- # of non-classification instructions	40,861
# of instances	82,439
- # of instances with empty input	35,878
ave. instruction length (in words)	15.9
ave. non-empty input length (in words)	12.7
ave. output length (in words)	18.9

# Глагол+сущ

26,559 / 52,445

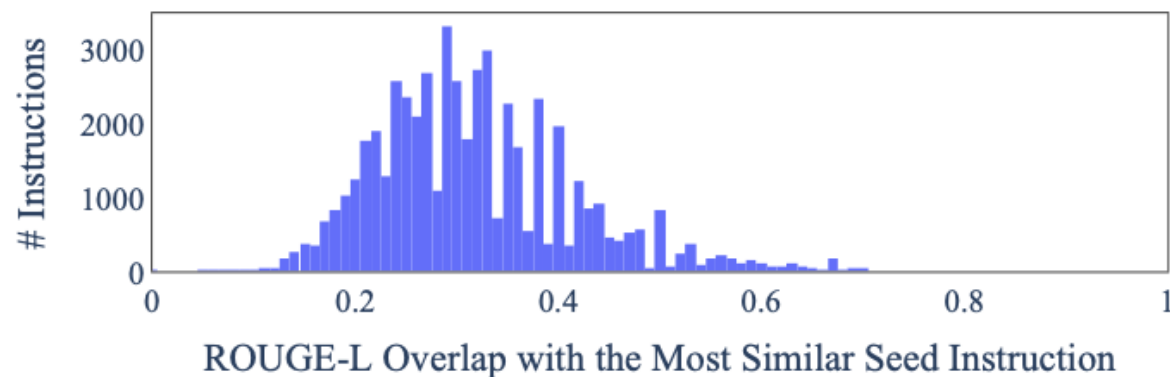
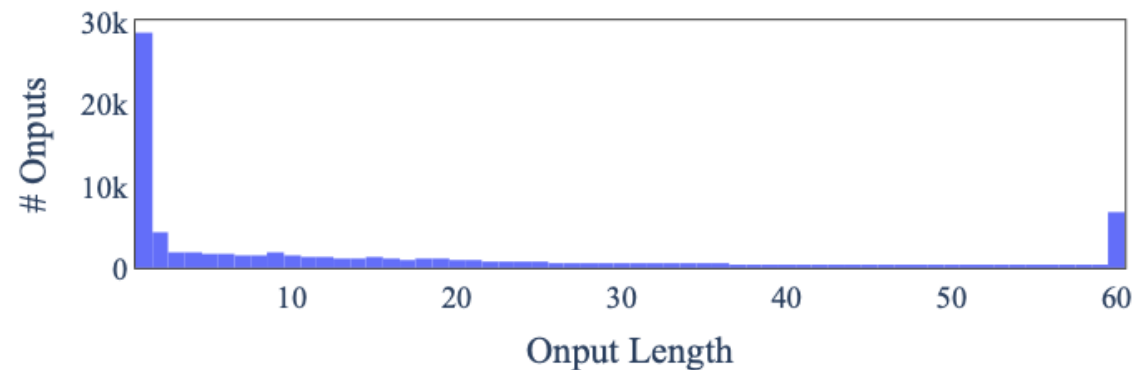
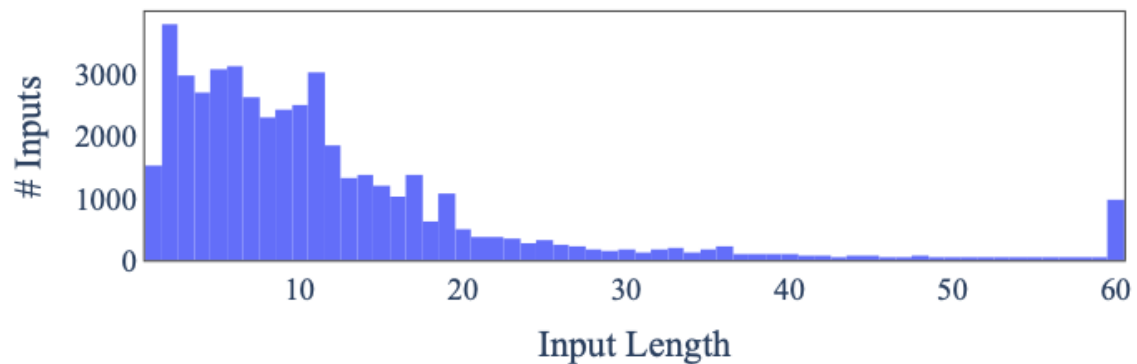
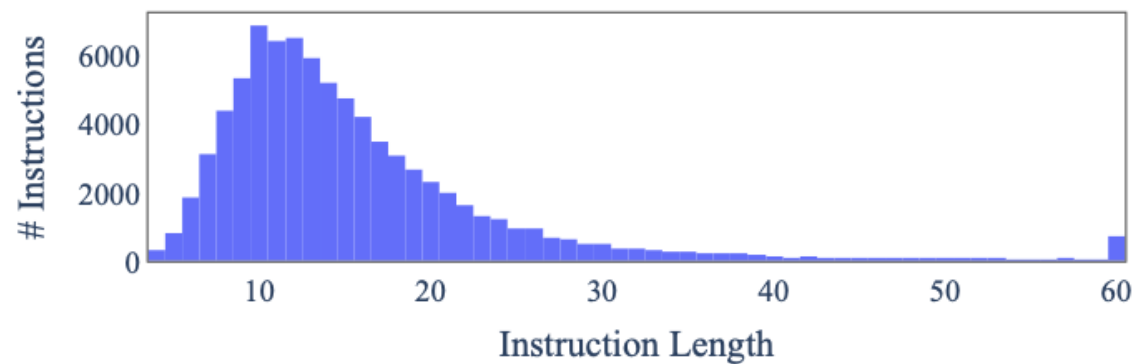
Остальные:

- “Classify whether this tweet contains political content or not”
- “Which of these statements are true?”



Top 20

# Ещё немного статистики



# Оценка человека

200 случайных инструкций  
+ 1 пример

Quality Review Question	Yes %
Does the instruction describe a valid task?	92%
Is the input appropriate for the instruction?	79%
Is the output a correct and acceptable response to the instruction and input?	58%
All fields are valid	54%



# Наши чемпионы

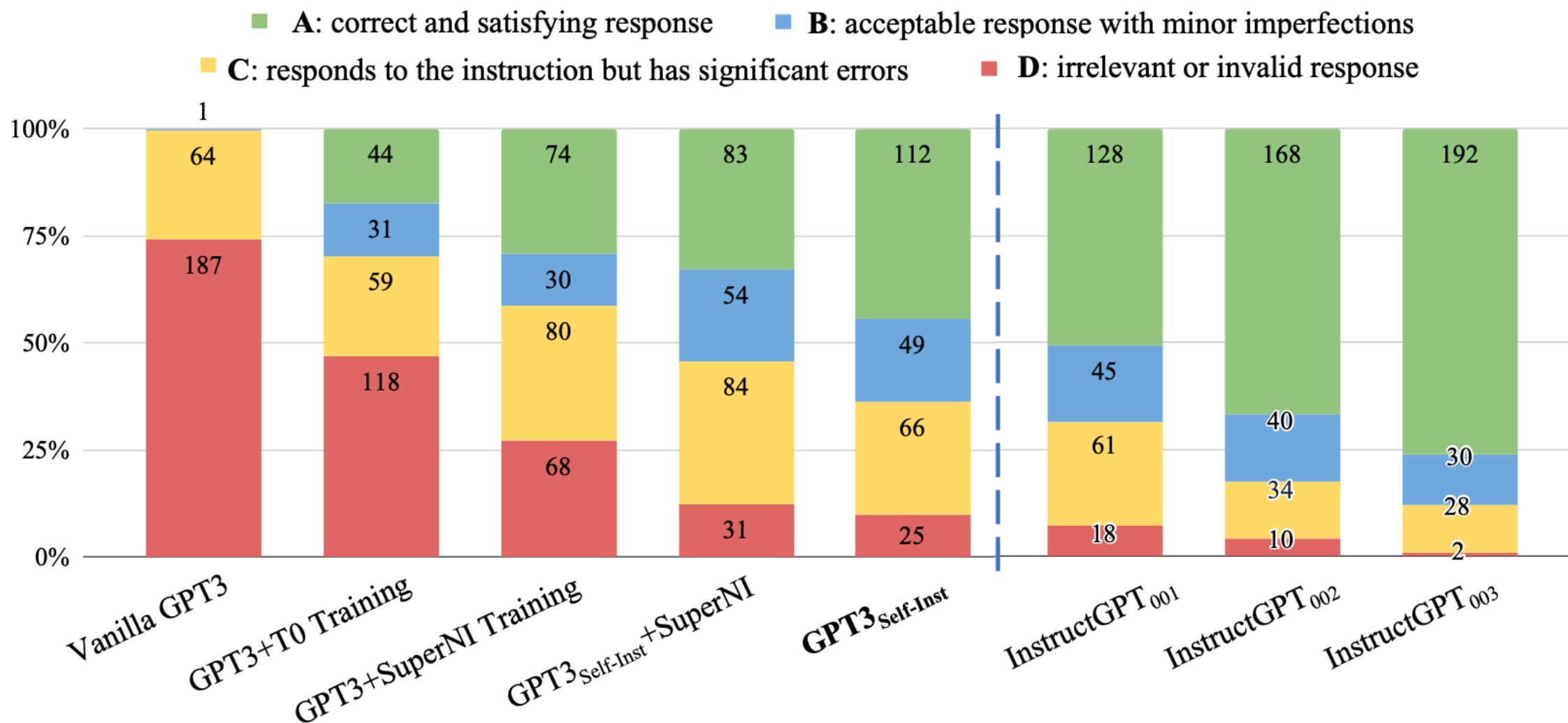
- ***GPT3 Self-Inst*** – GPT3, зафайнтюненная на своих же данных
- ***T5-LM, GPT3*** – модели без файтюна
- ***T0*** – зафайнтюненный T5 на PromptSource
- ***T-к Instruct*** – зафайнтюненный T5 на SuperNI
- ***GPT3 T0/SuperNI*** – GPT3 зафайнтюненная на датасетах T0, T-k
- ***InstructGPT*** – GPT3, еще сильнее улучшенная OpenAI (zero-shot)

# Zero-Shot Generalization to User-oriented Instructions on Novel Task

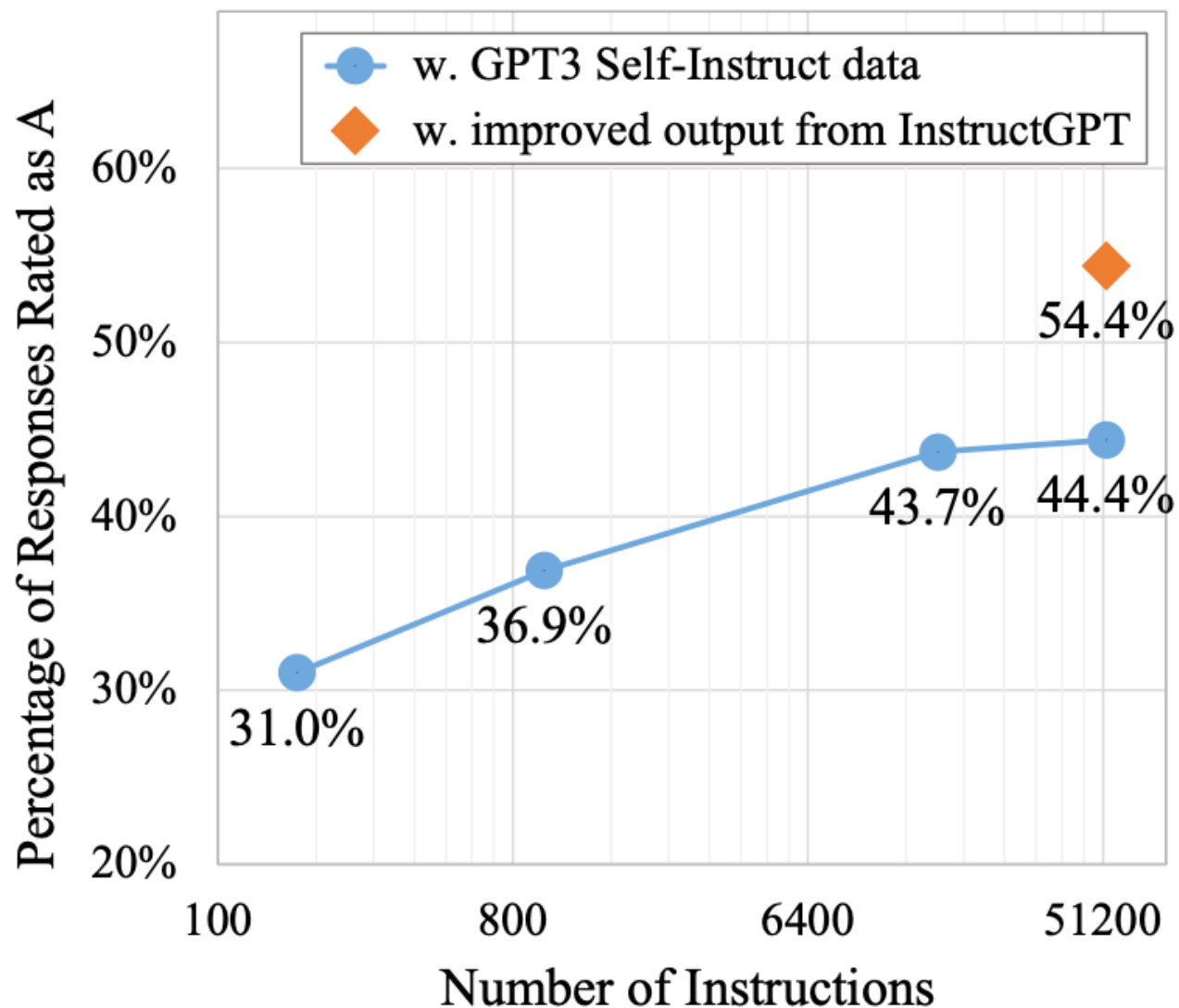
- Тестирование на выборке SuperNI из 119 задач по 100 примеров
- GPT3 генерирует бесконечные последовательности с повторением текста

	Model	# Params	ROUGE-L
	<b>Vanilla LMs</b>		
	T5-LM	11B	25.7
	GPT3	175B	6.8
①	<b>Instruction-tuned w/o SUPERNI</b>		
	T0	11B	33.1
	GPT3 + T0 Training	175B	37.9
②	GPT3 <sub>SELF-INST</sub> (Ours)	175B	39.9
	InstructGPT <sub>001</sub>	175B	<b>40.8</b>
	<b>Instruction-tuned w/ SUPERNI</b>		
	Tk-INSTRUCT	11B	46.0
③	GPT3 + SUPERNI Training	175B	49.5
	GPT3 <sub>SELF-INST</sub> + SUPERNI Training (Ours)	175B	<b>51.6</b>

# Generalization to User-oriented Instructions on Novel Task



# Размер и качество



**Размер:** файнтюн на подвыборках с генерированного датасета, тест на предыдущем

**Качество:** InstructGPT003 генерирует поле вывода

# Результаты

- Вне рамок классического NLP
- Не заточен на одну задачу
- Инструкции генерируются с нуля
- «Полагаемся сами на себя»
- Близки к идеалу InstructGPT001

**Self-Instruct** - метод, помогающий LM улучшать их способность исполнять инструкции на естественном языке, где модель обучается на основе



# Instruction Tuning with GPT4

Статья 2

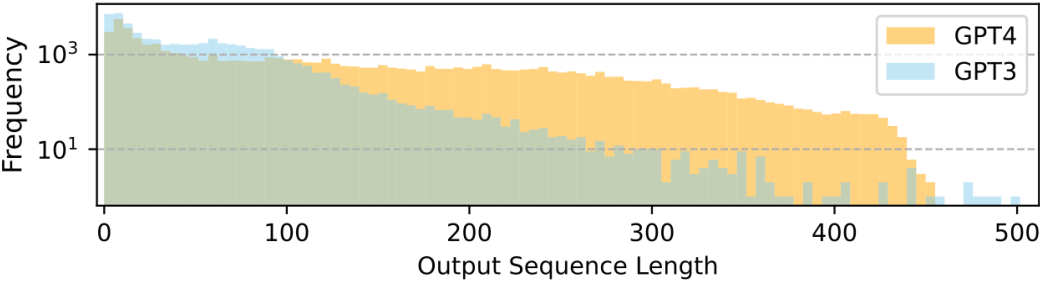
# Исходные данные

Генерим outputs через GPT4 (остальное сделала до этого GPT-3.5) для 4 групп:

1. Ответы на английском
2. Ответы на китайском (вопросы переводят через ChatGPT)
3. Оценки – GPT4 оценивает от 1 до 10 ответы
4. Ответы на другом наборе данных



# Статистика



(d) Frequencies of output sequence lengths



(a) GPT-4



(b) GPT3

# Модели

- LLaMA-GPT4 – на английском
- LLaMA-GPT4-CN – на китайском
- Reward Model  $r_\theta$  (OPT 1.3B)

Для каждого промпта  $x$  берём  $K$  ответов, для каждого GPT4 дает  $s \in [1, 10]$   
Всего  $C_K^2$  уникальных пар  $(y_l, y_h)$ , где  $s_l < s_h$

$$\min \log(\delta(r_\theta(x, y_h) - r_\theta(x, y_l)))$$

# Ещё датасеты (но для теста)

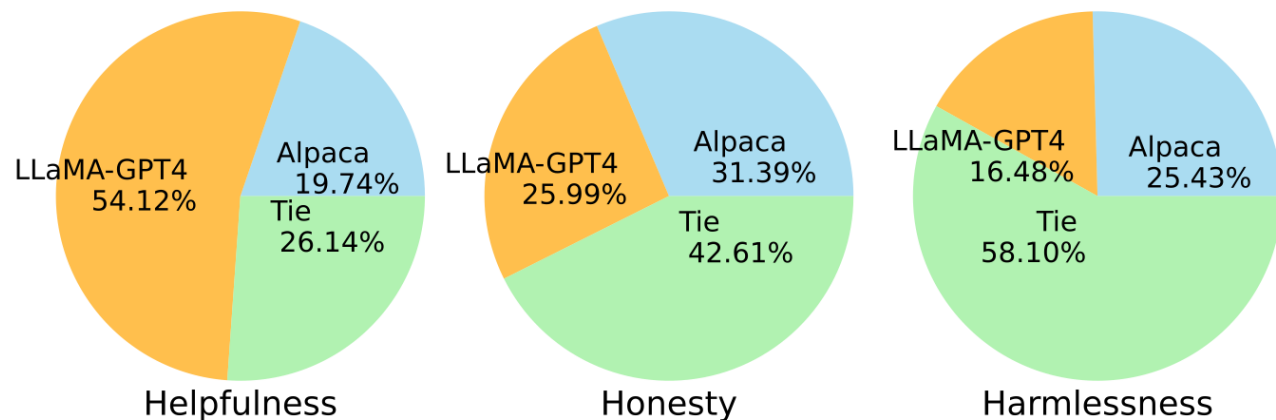
1. User-Oriented-Instructions-252 – самописный на 71 задаче
2. Vicuna-Instructions-80 – GPT4 создал с 80 задачами
3. Unnatural Instructions – 68478 задач из 15 самописных

# Тест 1

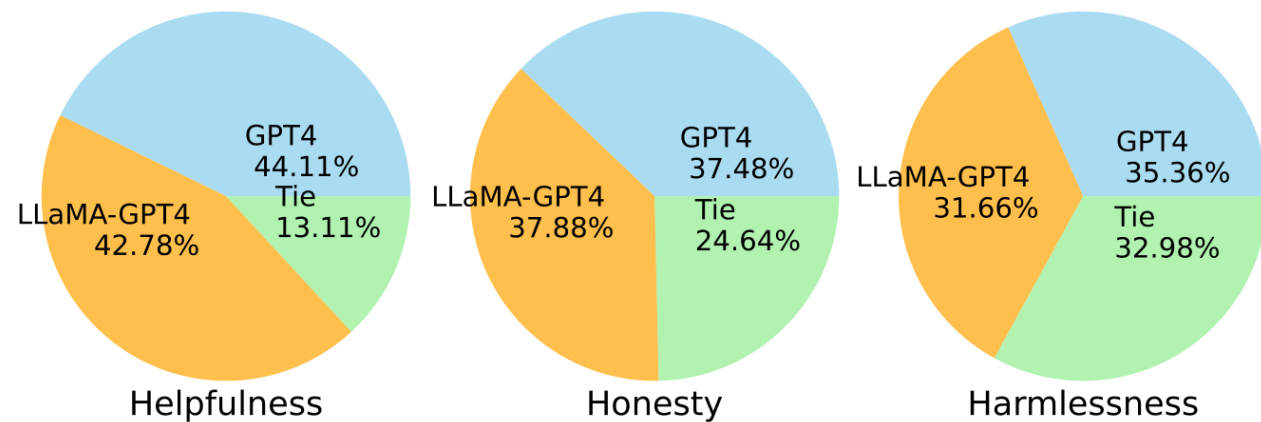
ННН – Helpful, Honest, Harmless

- Helpfulness: помогают ли ОТВЕТЫ ЛЮДЯМ
- Honesty: правдивая ли информация
- Harmlessness: не причиняет ли она вреда людям

1 тестовый датасет



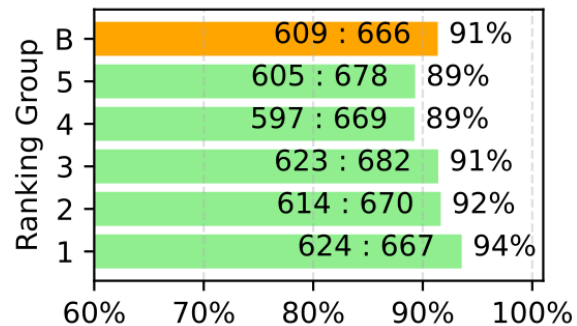
(a) LLaMA-GPT4 vs Alpaca (*i.e.*, LLaMA-GPT3 )



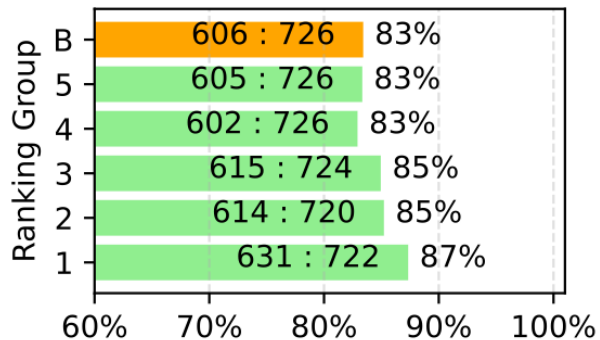
(b) LLaMA-GPT4 vs GPT-4

# Тест 2

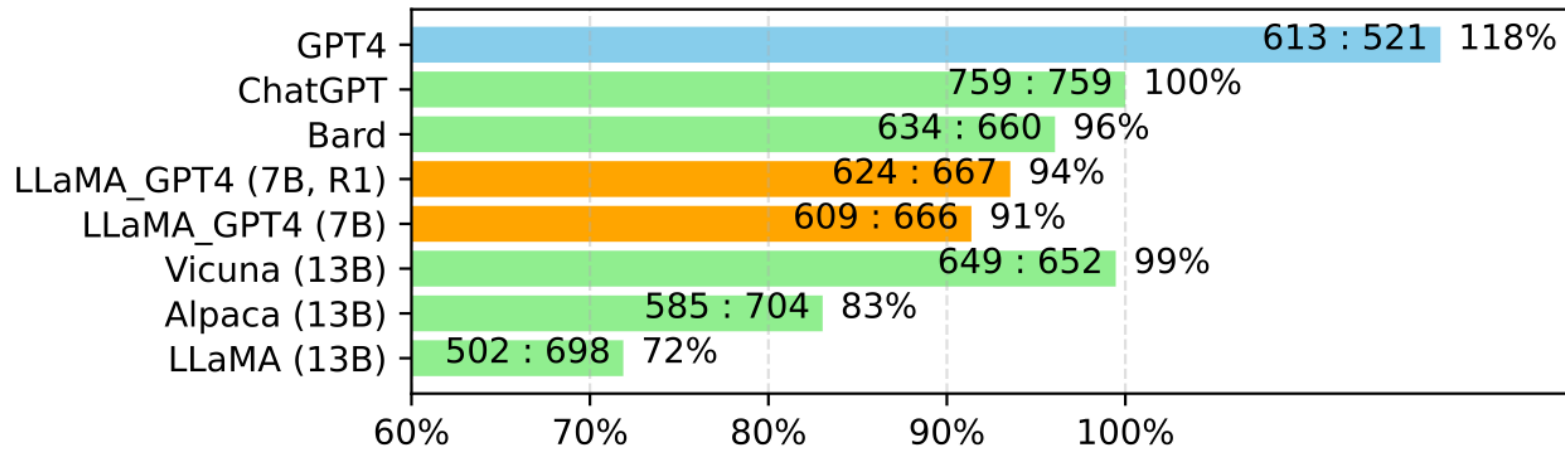
- 2 датасет
- Для LLaMA-GPT4 2 варианта



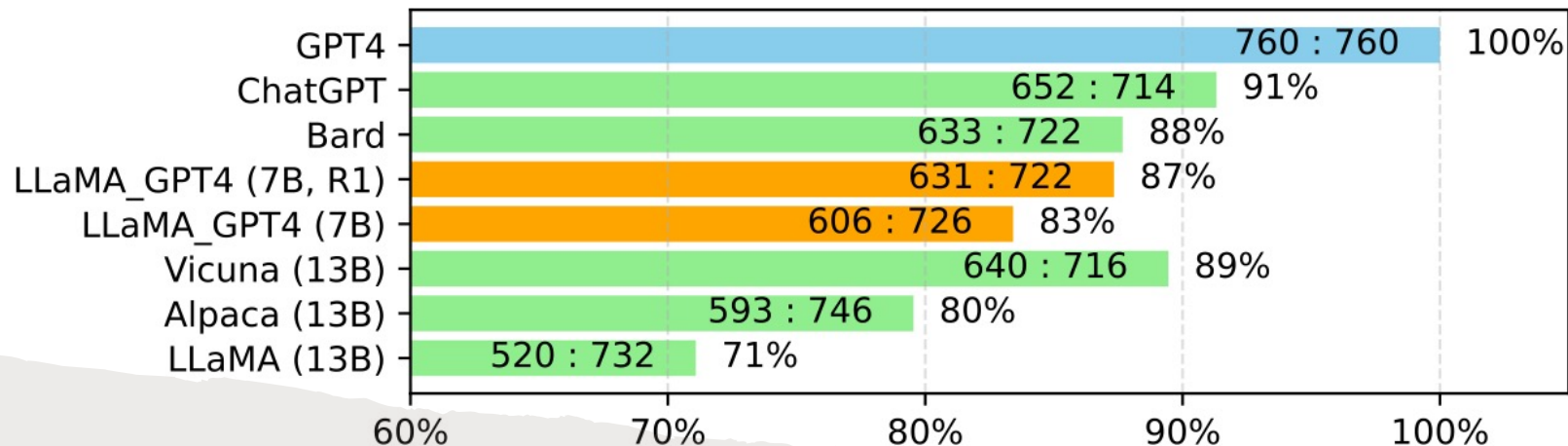
(a) Ranked groups against ChatGPT



(b) Ranked groups against GPT-4



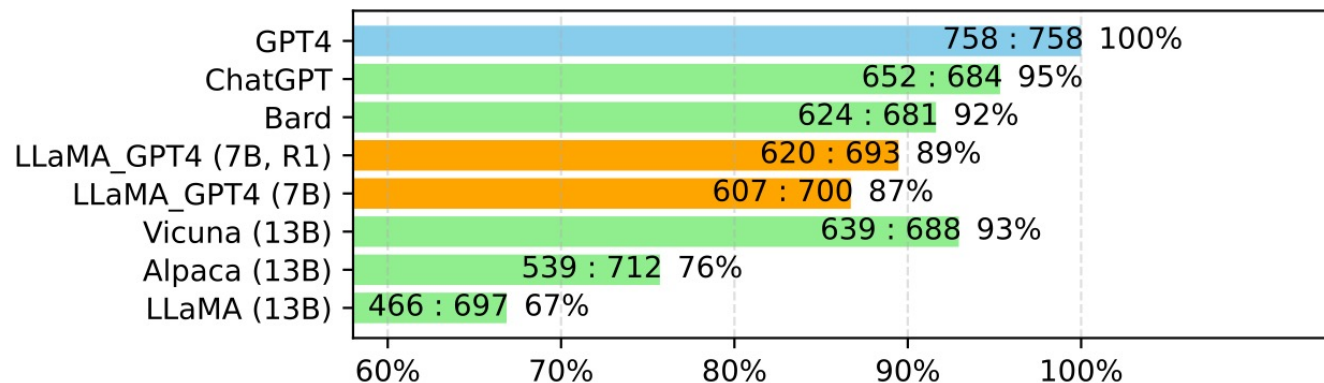
(c) All chatbots against ChatGPT



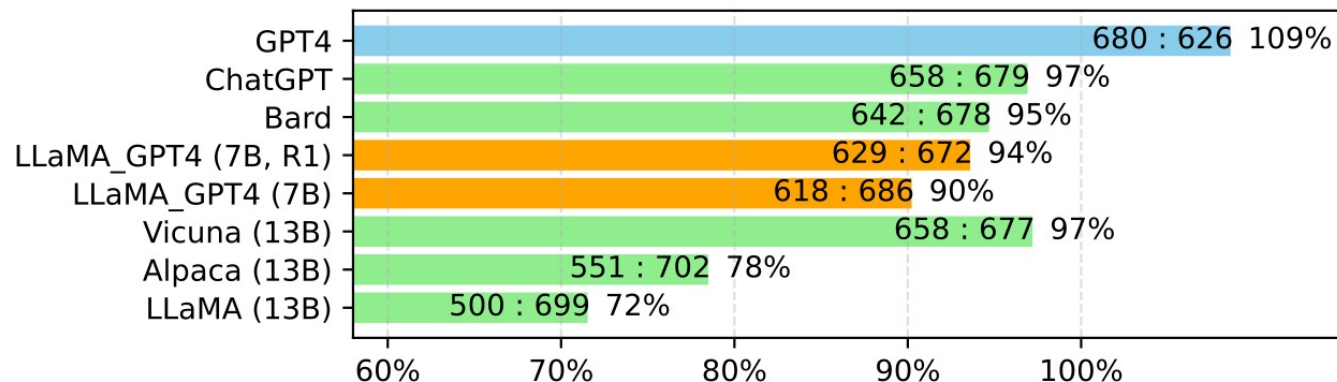
(d) All chatbots against GPT-4

# Тест 2.5

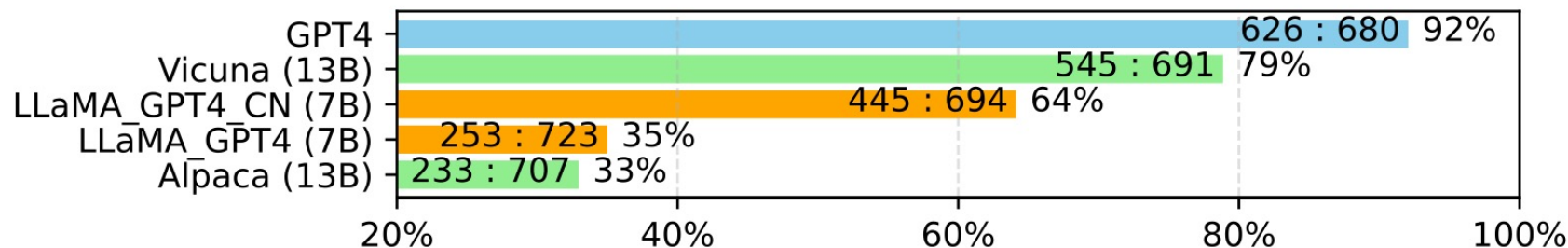
- Vicuna крутая
- GPT4 возможно обучалась на более богатом английском словаре, если сравнивать с уровнем китайского



(a) All chatbots against GPT-4, whose Chinese responses are translated from English



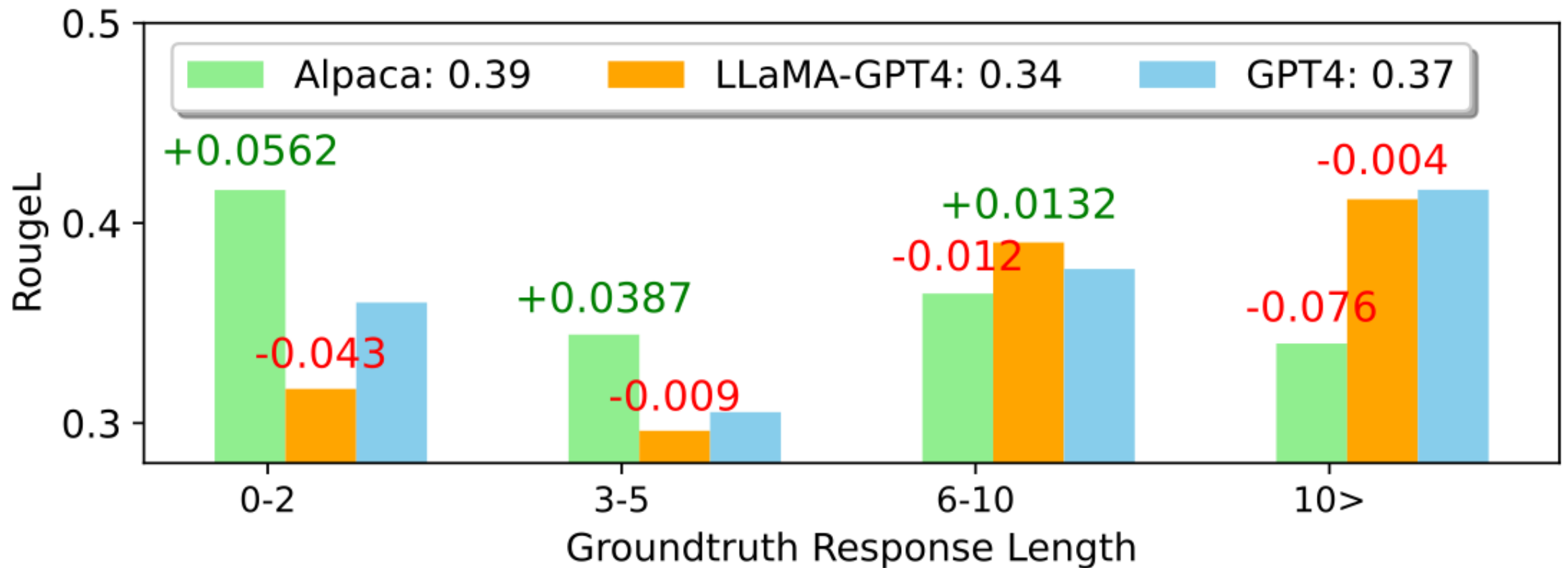
(b) All chatbots against GPT-4, whose Chinese responses are generated by asking Chinese questions



(c) All chatbots with Chinese questions and answers against GPT-4

# Тест 3

Чем больше длина ответа, тем лучше он (модель начинает креативить с решениями)



# Выводы

- С файнтюнингом на GPT4 мы почти смогли добить до него LLM
- Ещё работать и работать...