

Генеративные разговорные модели: метрики, данные, перспективы

Валерий Архипов
R&D NLP
@SberDevices

SBER

DEVICES

Сегодня в программе:

- про chitchat-модели в ассистентах Сбера
- про архитектуру и обучение базовой языковой модели (LM)
- про переход от обычной LM к диалоговой LM
- про трюки предобучения, дообучения и особенности датасетов
- про метрики диалоговых моделей
- покажем интересные примеры генерации

Сегодня в программе (2):

- Выводы по проделанной работе:

постпроцессинг, ускорение, очистка данных, вклад датасетов

- Что дальше?

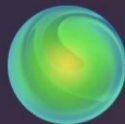
Обогащение датасетов, RL для читчата (!)

Ассистенты «Салют!»

3 персонажа, разные диалоговые системы

Сбер

Деловой стиль общения как у сотрудника Сбера



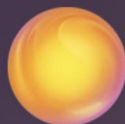
Афина

Умеренный тон, понимающий собеседник для любых задач



Джой

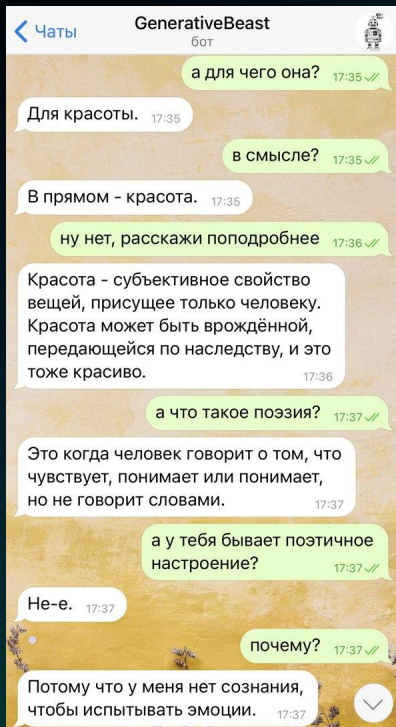
Лёгкий стиль общения и бодрое настроение



- **Retrieval** — диалоговая модель для Сбера и Афины

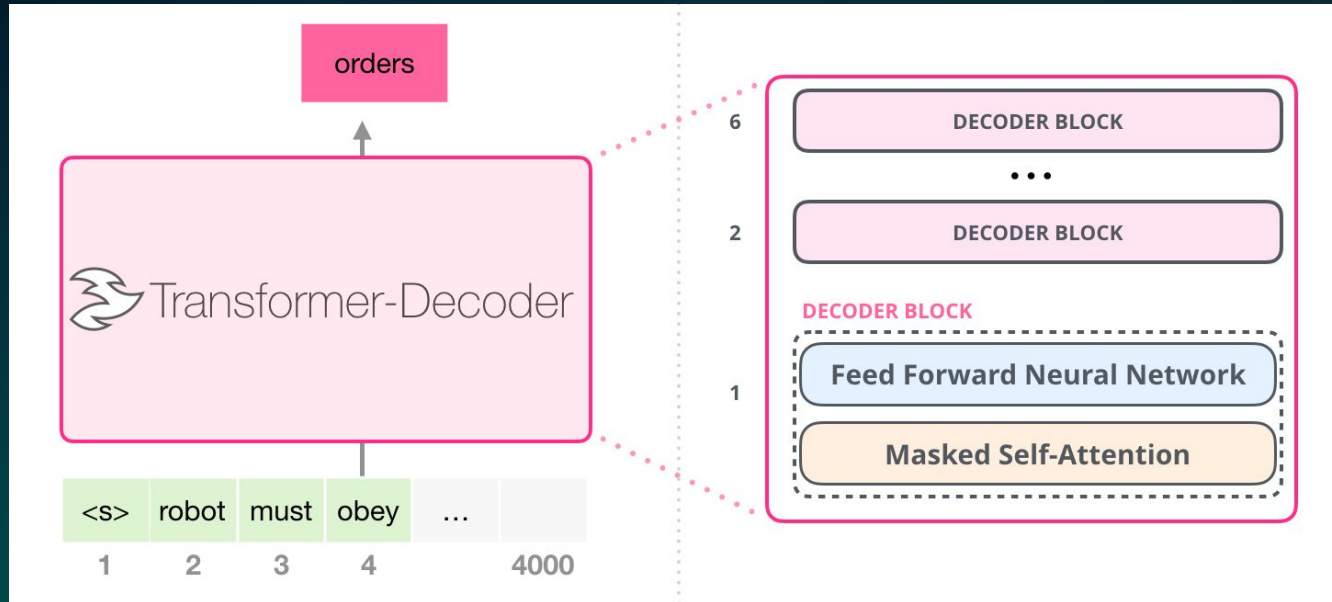
- **Генеративная модель** для Джой

Немного примеров



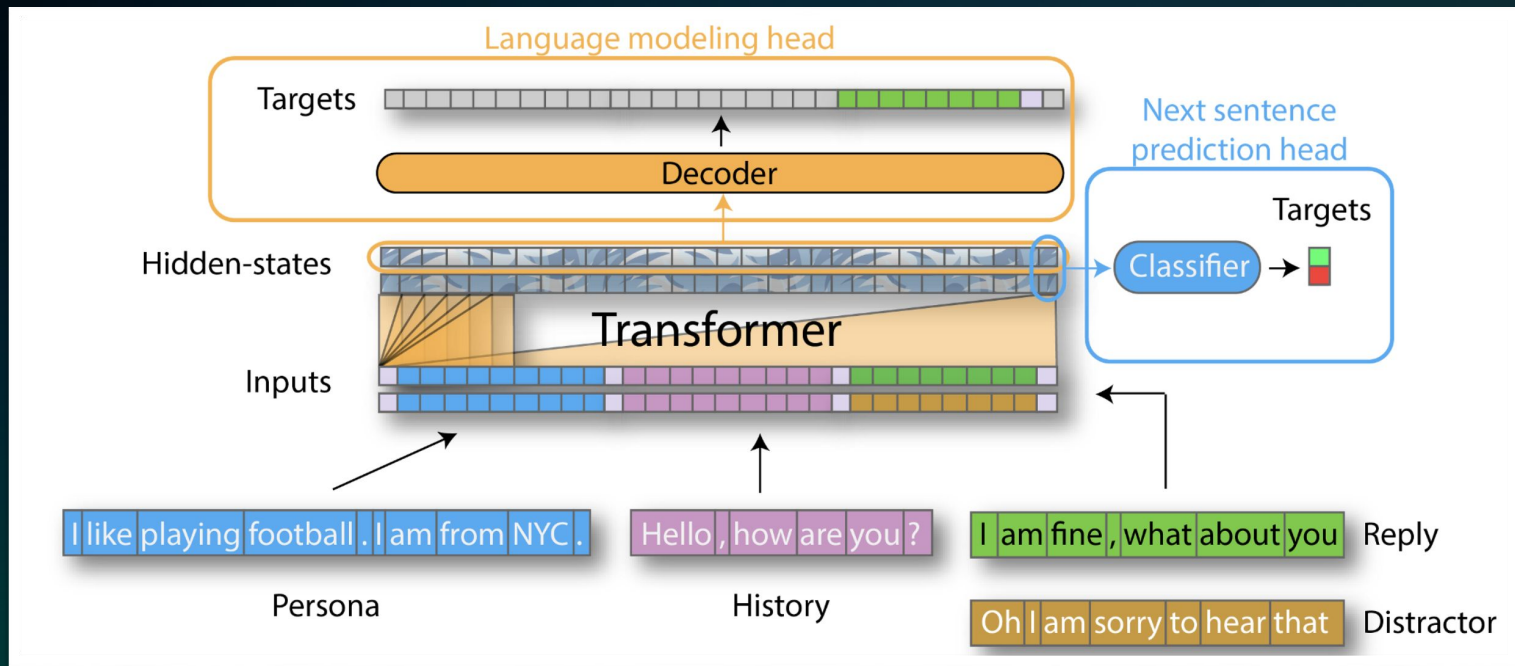
Генеративная модель GPT-2 (3)

Аutoreгрессионная decoder-only модель, наблюдает контекст слева для предсказания следующего токена.



Переход к диалоговой модели

- Модификация архитектуры GPT-2 (ConvAI Hugging Face)



Переход к диалоговой модели

- **Специальные датасеты**

Для обучения качественной диалоговой модели жизненно необходим набор диалогов, максимально удовлетворяющий следующим свойствам:

- стилистическая схожесть ответов
- широкий набор тем
- соблюдение контекста диалога
- ответы интересные и нетривиальные

Pre-training → Fine-tuning

Parent LM:

- только LM голова нейросети
- 180 Гб RU данных (книги, вики, новости)
- Adam, $lr = 0,00015$, linear decay
- ≈ 5 epochs
- $\approx 200k$ steps
- perplexity ≈ 12 (valid)

Dialogue LM:

- 2 головы: LM + NSP
- 0,5 Гб диалогов (вместе с кандидатами для NSP)
- NovoGrad, $lr = 0,015$, linear decay
- 20 epochs
- $\approx 10k$ steps
- perplexity ≈ 5 (dialogs set)

Pre-training → Fine-tuning

Эмпирический подбор параметров:

- `lm_loss` vs. `mc_loss` (2 к 1)
- Число кандидатов для NSP (2—7)
- Negative Sampling (перешли от random к LSTM к catboost)
- Длина контекста (глобальное ограничение в 1024 токена (как у родительской модели), но не более 128 токенов на реплику у диалоговой, смотрим на последние 8 реплик)

Генерация ответов диалоговой модели — ranking

Генерируем несколько (≤ 10) кандидатов ответов с помощью Nucleus sampling, после чего ранжировать:

- LSTM (64)
- Catboost
- NSP head
- Простыми эвристиками

Такой трюк позволил увеличить SSA на ≈ 10 п. п.

Генерация ответов диалоговой модели

Nucleus Sampling

- $\text{top_p} = 0,8—0,95$ (отбор наиболее вероятных токенов)
- $\text{temperature} = 0,8—1,2$ (повышением снижаем уверенность модели)

Либо **жадная генерация** каждого токена (top-1)

жадная генерация

>>> кто такой Ленин?
> Ленин - российский политический и
государственный деятель, лидер
Коммунистической партии Советского Союза

nucleus

>>> кто такой Ленин?
> Есть много разных версий биографии,
некоторые более-менее достоверные

Метрики

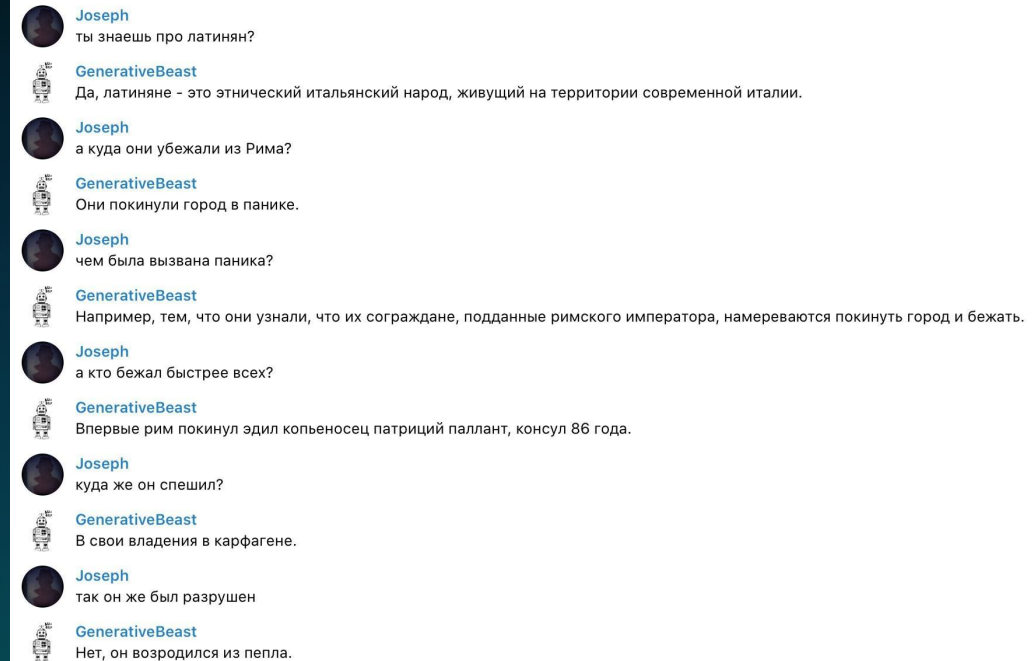
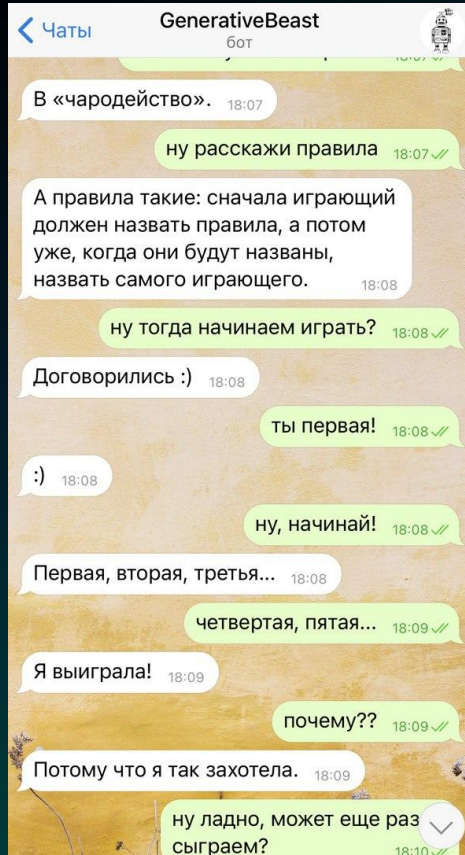
SSA by Google Meena (Sensibility Specificity Average)

- Оценка абсолютная, без проблем могут сравниваться модели разных поколений между собой
- Оценка относительно трудная для разметчиков
- Интерпретация результатов не всегда очевидная

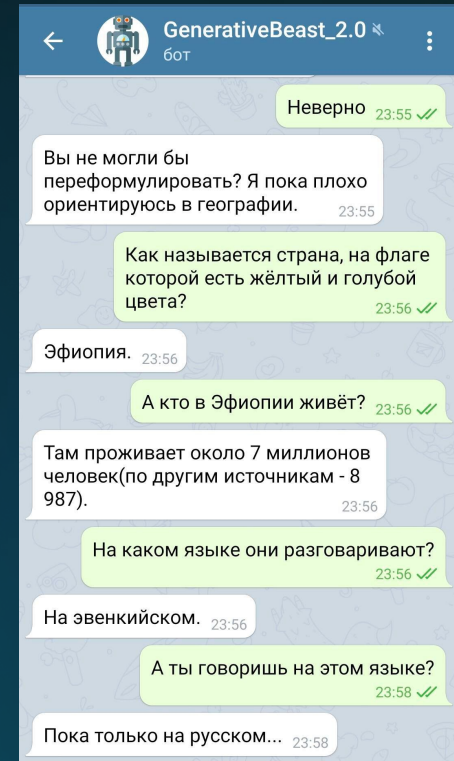
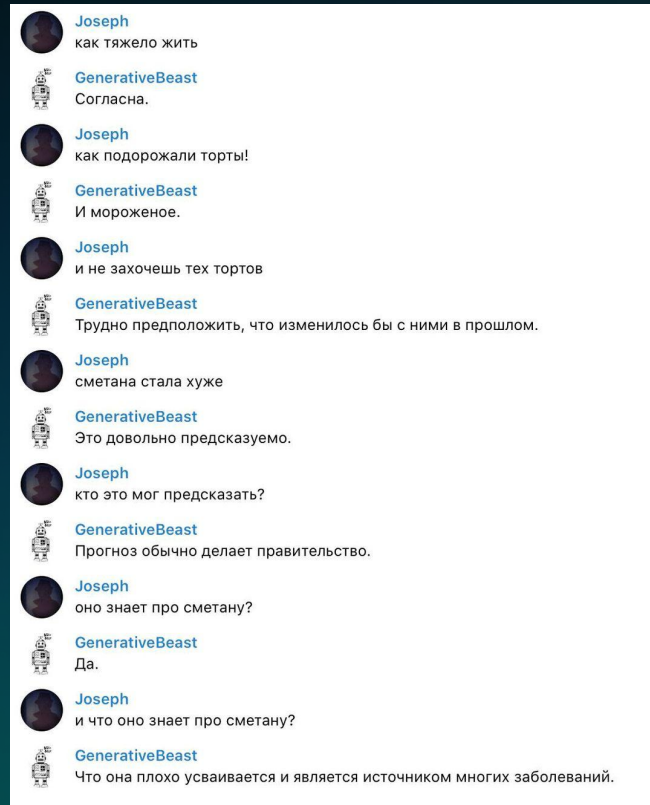
ACUTE-eval by Facebook BlenderBot

- Оценка относительная, попарные сравнения
- Простая разметка
- Можно использовать Bot-Bot диалоги для прокси-оценки
- Интерпретация результатов с точки зрения продукта очевидна

примеры генерации



примеры генерации



Выводы по проделанной работе

Сборник советов

- с определенного момента единый стиль и качество ответов важнее, чем количество диалогов
- постобработка - боль и зло
- сухие фактоидные “диалоги” не всегда улучшают e2e качество
- используйте батчинг, ограничение длины генерации, ограничение длины контекста, кэш, fp16, мЕньшие модели, onnxruntime
- следите за разметчиками, оценивающими реплики для метрик
- Играйте с параметрами генерации - topp, topk, temp, rep_penalty, {beam search}
- Ранжируйте кандидатов

Что дальше — Offline reinforcement learning

Что хотим?

- улучшать генерацию
- улучшать постпроцессинг
- использовать постоянно увеличивающуюся базу логов

Какие вопросы?

- как расширять редакторские сетки понятно, как использовать реальные диалоги с пользователями не очень
- как именно дообучать модель на этом?
- будет ли реальный позитивный эффект?

Что дальше — Offline reinforcement learning

Human-centric dialog training via offline reinforcement learning

**Natasha Jaques^{*12}, Judy Hanwen Shen^{*1}, Asma Ghandeharioun¹, Craig Ferguson¹,
Agata Lapedriza¹, Noah Jones¹, Shixiang Shane Gu², Rosalind Picard¹**

***Equal contribution**

¹Massachusetts Institute of Technology, Cambridge, USA

`<judyshen, asma_gh, agata, ncjones, roz>@mit.edu`

²Google Research, Mountain View, USA

`<natashajaques, shanegu>@google.com`

Что дальше — Offline reinforcement learning

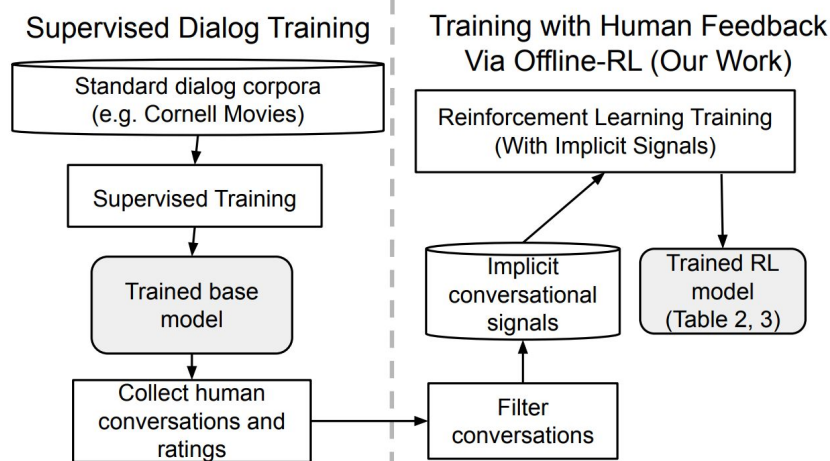


Figure 1: Schematic diagram of our method for training with human conversation cues via offline RL. Unlike traditional approaches which stop at using explicit feedback to evaluate static conversations, we allow humans to freely interact with dialog models, and compute metrics based on their implicit satisfaction which are optimized using offline RL.

$$\Psi^*(s_t, a_t) = r(s_{t'}, a_{t'})/c + \log p(a_{t'}|s_{t'}) + \gamma \log \sum_{a'} \exp(\Psi^*(s', a')) \quad (4)$$

$$\pi_{\Psi}^*(a_t|s_t) = \exp(\Psi^*(s_t, a_t)) \quad (5)$$

Зачем надевать наволочку
на подушку?

Чтобы она не промокла от
соплей

Спасибо!

 SBER DEVICES



Валерий Архипов

давай ржачный анекдот про Штирлица и Мюллера!



GenerativeBeast_2.0

Штирлиц был хорошим психологом и заставил Мюллера признаться в шпионаже