

Тест начат понедельник, 11 марта 2024, 12:23

Состояние Завершены

Завершен понедельник, 11 марта 2024, 12:54

Прошло времени 31 мин. 13 сек.

Оценка 5,00 из 5,00 (100%)

Вопрос 1

Выполнен

Баллов: 1,00 из 1,00

Алгоритм машинного перевода переводит текст с английского языка на русский. Размер английского словаря равен 5000 токенов, русского - 8000. Оцените, сколько последовательностей рассмотрит декодер при генерации перевода наибольшей длины 5 при условии, что при декодировании используется **алгоритм полного перебора**?

В ответ запишите одно число.

Подсказка: в результате получится достаточно большое число.

Ответ: 3,2768e+19

Ваш ответ верный.

Вопрос 2

Выполнен

Баллов: 1,00 из 1,00

Алгоритм машинного перевода работает с переводом текста с английского языка на русский. Размер английского словаря равен 5000 токенов, русского - 8000. Оцените, сколько последовательностей сгенерирует декодер при переводе предложения длины 5 при условии, что используется **лучевой поиск (beam search)** при $b=3$ (b - ширина луча)?

В ответ запишите одно число без пробелов.

Ответ: 243

Ваш ответ верный.

Вопрос 3

Выполнен

Баллов: 1,00 из 1,00

Есть ли различие между жадным поиском и лучевым поиском (beam search) с $k=1$?

Размер словаря равен $|V|$, длина последовательности для перевода равна n .

Выберите один ответ:

- ☐ a. Есть, жадный поиск сгенерирует один вариант перевода, а лучевой поиск n
- ☐ b. Есть, при жадном поиске алгоритм рассмотрит все возможные варианты перевода ($|V|^n$)
- ☒ c. Нет

Ваш ответ верный.

Вопрос 4

Выполнен

Баллов: 1,00 из 1,00

Пусть у нас есть предложение, состоящее из произвольных цифр. Например, '1576429830'.

Посчитайте перплексию этого предложения относительно модели, которая считает, что вероятность встретить каждую цифру равна 0.1.

*Напоминание: $PP(W) = P(w_1 w_2 \dots w_N)^{-1/N}$

Ответ:

Ваш ответ верный.

Вопрос 5

Выполнен

Баллов: 1,00 из 1,00

BLEU (bilingual evaluation understudy) - метрика для оценивания качества машинного перевода, основанная на сравнении перевода, предложенного алгоритмом, и референсного перевода (ground truth). Сравнение производится на основе подсчета n-грамм (n меняется от 1 до некоторого порога, например, 4), которые встретились и в предложенном переводе, и в референсном (ground truth).

После подсчета совстречаемости n-грамм полученная метрика умножается на так называемый brevity penalty - штраф за слишком короткие варианты перевода. Brevity penalty считается как <количество слов в переводе, предложенном алгоритмом> / <количество слов в референсном переводе>.

Формула:

$$BLEU = \text{brevity penalty} \cdot (\prod_{i=1}^n \text{precision}_i)^{1/n} \cdot 100\%, \text{ где } \text{brevity penalty} = \min \left(1, \frac{\text{output length}}{\text{reference length}} \right)$$

Пример:

SYSTEM A: Israeli officials responsibility of airport safety
2-GRAM MATCH 1-GRAM MATCH

REFERENCE: Israeli officials are responsible for airport security

SYSTEM B: airport security Israeli officials are responsible
2-GRAM MATCH 4-GRAM MATCH

Metric	System A	System B
precision (1gram)	3/6	6/6
precision (2gram)	1/5	4/5
precision (3gram)	0/4	2/4
precision (4gram)	0/3	1/3
brevity penalty	6/7	6/7
BLEU	0%	52%

Задача

Посчитайте BLEU-score для следующего предложения. При подсчете метрики учитывайте n-граммы с $n \in [1, 2, 3]$.

Перевод, предложенный алгоритмом: "Кошка вышла из дома и села на крыльцо"

Референсный перевод (ground truth): "Кошка вышла из комнаты и села на ступеньки"

Формат ответа: ответ запишите в виде процентов, округлив до целых.

Ответ:

Ваш ответ верный.

◀ 5.5 Семинар: генерация кода со Stack Overflow

Перейти на...

6.1 Контекстуализированные представления и перенос знаний ▶