

Применение больших языковых моделей к решению математических задач

Изуткина А.М. ML12-22

Московский Государственный Технический Университет им. Н.Э.Баумана
Кафедра ФН12 «Математическое моделирование»

2024г

Постановка задачи

Цель работы: Изучить принципы и подходы, лежащие в основе больших языковых моделей (LLM), с их помощью написать программу генерации решений математических задач.

- 1 Изучить основные подходы и архитектуры LLM в данной области.
- 2 Изучить подходы дообучения (Fine-Tuning).
- 3 Реализовать класс датасета, эффективный цикл обучения, валидацию.
- 4 Оценить результаты.

QKV Attention

Создание матриц Q, K и V: Каждый входной вектор преобразуется в три различных вектора с помощью обучаемых матриц W_Q , W_K , W_V :

$$\mathbf{Q} = \mathbf{X}W_Q,$$

$$\mathbf{K} = \mathbf{X}W_K,$$

$$\mathbf{V} = \mathbf{X}W_V.$$

Вычисление оценок:

$$\text{scores} = \frac{\mathbf{Q} \cdot \mathbf{K}^T}{\sqrt{d_k}} \quad (1)$$

где $\mathbf{Q} \in \mathbb{R}^{n \times d_k}$, $\mathbf{K} \in \mathbb{R}^{n \times d_k}$, $\mathbf{V} \in \mathbb{R}^{n \times d_v}$.

Применение софтмакса для нормализации:

$$\text{attention weights} = \text{softmax}(\text{scores}), \quad (2)$$

Вычисление контекстного вектора:

$$\text{output} = \text{attention weights} \cdot \mathbf{V} \quad (3)$$

MultiHead Attention

1. Каждая голова выполняет свой собственный механизм самовнимания:

$$\text{head}_i = \text{Attention}(Q_i, K_i, V_i)$$

2. Результаты всех голов конкатенируются:

$$Z = \text{Concat}(\text{head}_1, \dots, \text{head}_h)$$

3. Итоговая матрица проходит через линейный слой для объединения информации:

$$\text{MultiHead}(Q, K, V) = ZW_O$$

где W_O — матрица весов для объединяющего слоя.

Основные компоненты трансформера

- Эмбединги: Каждое слово преобразуется в векторное представление, что позволяет модели захватывать семантические отношения.
- Механизм внимания: Позволяет модели фокусироваться на различных частях входного текста, учитывая контекст при генерации следующего слова.
- Обучение с учителем: Модель обучается на больших объемах текстовых данных, что позволяет ей предсказывать вероятные последовательности слов.

Кросс-энтропийная функция потерь:

$$L(y, \hat{y}) = - \sum_{i=1}^C y_i \log(\hat{y}_i) \quad (4)$$

где $y_i \in \{0, 1\}$ — истинные метки, \hat{y}_i — предсказанные вероятности, а C — количество классов.

Parameter-Efficient Fine-Tuning (LoRA)

PEFT позволяет адаптировать большие языковые модели, минимизируя количество настраиваемых параметров.

В методе LoRA обновления весов представляются в виде двух матриц низкого ранга A и B :

Исходные веса модели W_0 остаются замороженными. Обновление весов представляется как:

$$W = W_0 + AB \tag{5}$$

где: $A \in \mathbb{R}^{d \times r}$, $B \in \mathbb{R}^{r \times d}$, $r < d$ — ранг, который определяет размерность обновления.

- Тип задач: Задачи охватывают широкий спектр тем, включая алгебру, геометрию и комбинаторику. В сумме 12 тыс. задач.
- Формат данных: Каждая запись в наборе данных содержит текст задачи и соответствующее решение, представленное в виде шагов.

MATH Dataset (Ours)

Problem: Tom has a red marble, a green marble, a blue marble, and three identical yellow marbles. How many different groups of two marbles can Tom choose?

Solution: There are two cases here: either Tom chooses two yellow marbles (1 result), or he chooses two marbles of different colors ($\binom{4}{2} = 6$ results). The total number of distinct pairs of marbles Tom can choose is $1 + 6 = \boxed{7}$.

Problem: If $\sum_{n=0}^{\infty} \cos^{2n} \theta = 5$, what is $\cos 2\theta$?

Solution: This geometric series is

$$1 + \cos^2 \theta + \cos^4 \theta + \dots = \frac{1}{1 - \cos^2 \theta} = 5. \text{ Hence,}$$

$$\cos^2 \theta = \frac{4}{5}. \text{ Then } \cos 2\theta = 2 \cos^2 \theta - 1 = \boxed{\frac{3}{5}}.$$

Problem: The equation $x^2 + 2x = i$ has two complex solutions. Determine the product of their real parts.

Solution: Complete the square by adding 1 to each side.

Then $(x + 1)^2 = 1 + i = e^{\frac{i\pi}{4}} \sqrt{2}$, so $x + 1 = \pm e^{\frac{i\pi}{8}} \sqrt[4]{2}$.

The desired product is then

$$\begin{aligned} & (-1 + \cos(\frac{\pi}{8}) \sqrt[4]{2}) (-1 - \cos(\frac{\pi}{8}) \sqrt[4]{2}) = \\ & 1 - \cos^2(\frac{\pi}{8}) \sqrt{2} = 1 - \frac{(1 + \cos(\frac{\pi}{4}))}{2} \sqrt{2} = \boxed{\frac{1 - \sqrt{2}}{2}}. \end{aligned}$$

Рис.: Пример из датасета MATH

1) Кросс-энтропия $H(p, q)$ между двумя распределениями p (истинное распределение) и q (предсказанное распределение) определяется как:

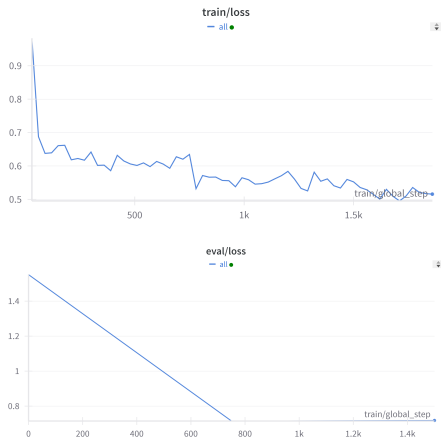
$$H(p, q) = - \sum_i p(i) \log(q(i))$$

$$2) \text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

3) BLEU - изначально разработана для оценки качества машинного перевода, но также широко используемая для оценки генерации текстов. Она измеряет степень совпадения между сгенерированным текстом и одним или несколькими эталонными текстами.

- Gemini 1.5 Pro от DeepMind
- o1-preview от OpenAI
- Qwen2.5-Math, DeepSeekMath и другие.

- 0.5 миллиарда обучаемых параметров.
- специально обучена для выполнения задач, связанных с пониманием и выполнением инструкций.
- Qwen2.5-Math показывает высокие результаты в решении математических задач.
- Использует ряд передовых технологий, включая:
 - RoPE (Rotary Positional Encoding): Позволяет улучшить обработку последовательностей с длинным контекстом.
 - SwiGLU: Функция активации, которая сочетает в себе преимущества различных функций активации для повышения производительности модели.
 - RMSNorm: Нормализация, которая помогает улучшить стабильность обучения.
 - Attention с учетом QKV Bias: Улучшает механизм внимания, позволяя модели более эффективно обрабатывать входные данные.



- В процессе дообучения Cross-entropy Loss упал с 1.6 до 0.7
- average BLEU: 0.2108
- part of correct answers: 0.2213

Instruction:

Please reason step by step, and put your final answer within $\boxed{}$

Input:

What $f(x)$ is a function such graph is $[-8, 2]$, and $g(x) = h\left(\frac{x^3}{2}\right)$, then what smallest of $g(x)$ is an interval of what length?

Response:

The f have $g(x) = h\left(\frac{x^3}{2}\right)$, the possible number x is in the domain of g if and only if $\frac{x^3}{2}$ is in the domain of h . Since The, the domain of g is of all numbers x such that $-8 \leq \frac{x^3}{2} \leq 8$. Multiplying interval to this inequality are x by $-16 \leq x \leq 16$, so the domain of g is x interval of length $16 - (-16) = \boxed{32}$. (

'True solution':

'Since we defined $g(x) = h\left(\frac{x^3}{2}\right)$, a real number x is in the domain of g if and only if $\frac{x^3}{2}$ is in the domain of h . Therefore, the domain of g consists of all x such that $-8 \leq \frac{x^3}{2} \leq 8$. The solutions to this inequality are given by $-16 \leq x \leq 16$, so the domain of g is an interval of width $16 - (-16) = \boxed{32}$.'

- Полезные библиотеки: unsloth.FastLanguageModel, trl.SFTTrainer

- Попробовать другие методы PEFT;
- Генерировать решения в виде python-кода;
- Для узко специализированных задач использовать методы RAG;