

高校数学単元を用いた類似問題推定による Few-shot プロンプトへの利用

田中 瞬十

十 広島市立大学院情報科学研究科

タイプ：数学タスク 参加枠：クローズ

はじめに

- Chain of Thought + 多数決 (Self-Consistency: SC) で回答を安定化
 - Few-shot プロンプトで回答精度を向上
 - 埋め込みによる類似問題推定では表面上を見るだけで回答の参考になる問題が推定できたか分からない
- ★単元ラベルによる類似問題推定手法を提案

チューニング手法

- Cot データ学習
- 複数回答を生成して多数決
- Few-shot プロンプト

埋め込みによる類似問題推定の課題

$\sqrt{x+2} = x$ を解け

単元ラベル

両辺を2乗して

$$x+2 = x^2 \text{ より } x^2 - x - 2 = 0$$

$$(x-2)(x+1) = 0$$

よって、 $x = 2, -1$

$$x \geq 0 \text{ より } x = 2$$

無理関数
不等式
因数分解
二次関数
二次方程式

$x+2 = x^2$ を解け

式を整理して

$$x^2 - x - 2 = 0$$

$$(x-2)(x+1) = 0$$

これより $x = 2, -1$

因数分解
二次関数
二次方程式

単元ラベルの定義

○単元ラベル (計72個)

平成30年告示学習指導要領数学編

単元ラベル：科目_大分類_中分類_小分類

○例

問： $y = x^2 + 2x + 3$ の最大値を求めよ

(数学 I_二次関数_値の変化_最大・最小)

$$\text{解：} x^2 + 2x + 3 = (x+1)^2 + 2$$

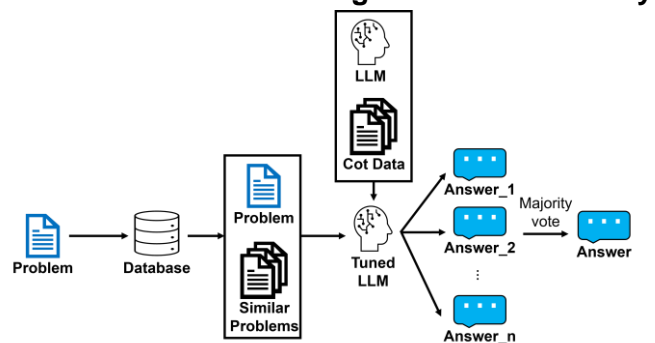
(数学 I_数と式_式_因数分解)

$y = (x+1)^2 + 2$ のグラフより最小値は2

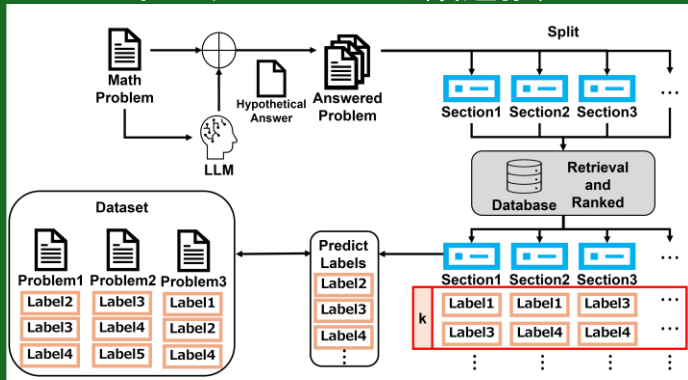
(数学 I_二次関数_そのグラフ_そのグラフ)

回答生成の流れ

Few-shot + Chain of Thought + Self-Consistency



単元ラベルによる類似問題推定



実験結果

評価問題数：100問

チューニングデータ、データセット：

Numinamath-CoT/synthetic_math の日本語訳19431問

データベース：Wikibooks

類似評価指標：Jaccard係数

LLM：v4-8b-instruct4

埋め込みモデル：Multilingual-e5-large

	中1	中2	中3	数I A	数II B	数III C	Overall
Baseline 10回平均	0.400	0.300	0.050	0.107	0.081	0.150	0.174
sc	0.600	0.389	0.063	0.133	0.192	0.300	0.270
SC+CoT	0.800	0.778	0.625	0.267	0.269	0.400	0.510
SC+CoT+ Fewshot (Embedding)	0.933	0.667	0.625	0.467	0.385	0.400	0.570
SC+CoT+ Fewshot (Unit Label) Top3	1.000	0.778	0.813	0.400	0.385	0.500	0.630