

## 研究背景

大規模言語モデル (Large Language Models, LLM) の進化

Transformer 構造を持つ BERT や  
Generative Pre-trained Transformer (GPT) など  
を活用した LLM の商業利用への需要拡大

⇒文章全体の分散表現を適切に得るための  
プーリング戦略が重要

## 先行研究

CLS-Average Pooling (CAP) 層の導入 [大和, 2024]<sup>[1]</sup>

学習可能な和が 1 となるパラメータ  $p, q (\geq 0)$   
を用いて  $E_{[CLS]}$  (文ベクトル項) 及び,  
 $E_{Avg}$  (平均単語ベクトル項)の重み付き和

$$pE_{[CLS]} + qE_{Avg}$$

を文の分散表現とする

⇒テキスト分類タスクにおいて,  $E_{[CLS]}$ ,  $E_{Avg}$  のみを  
用いた場合よりも高い性能を発揮

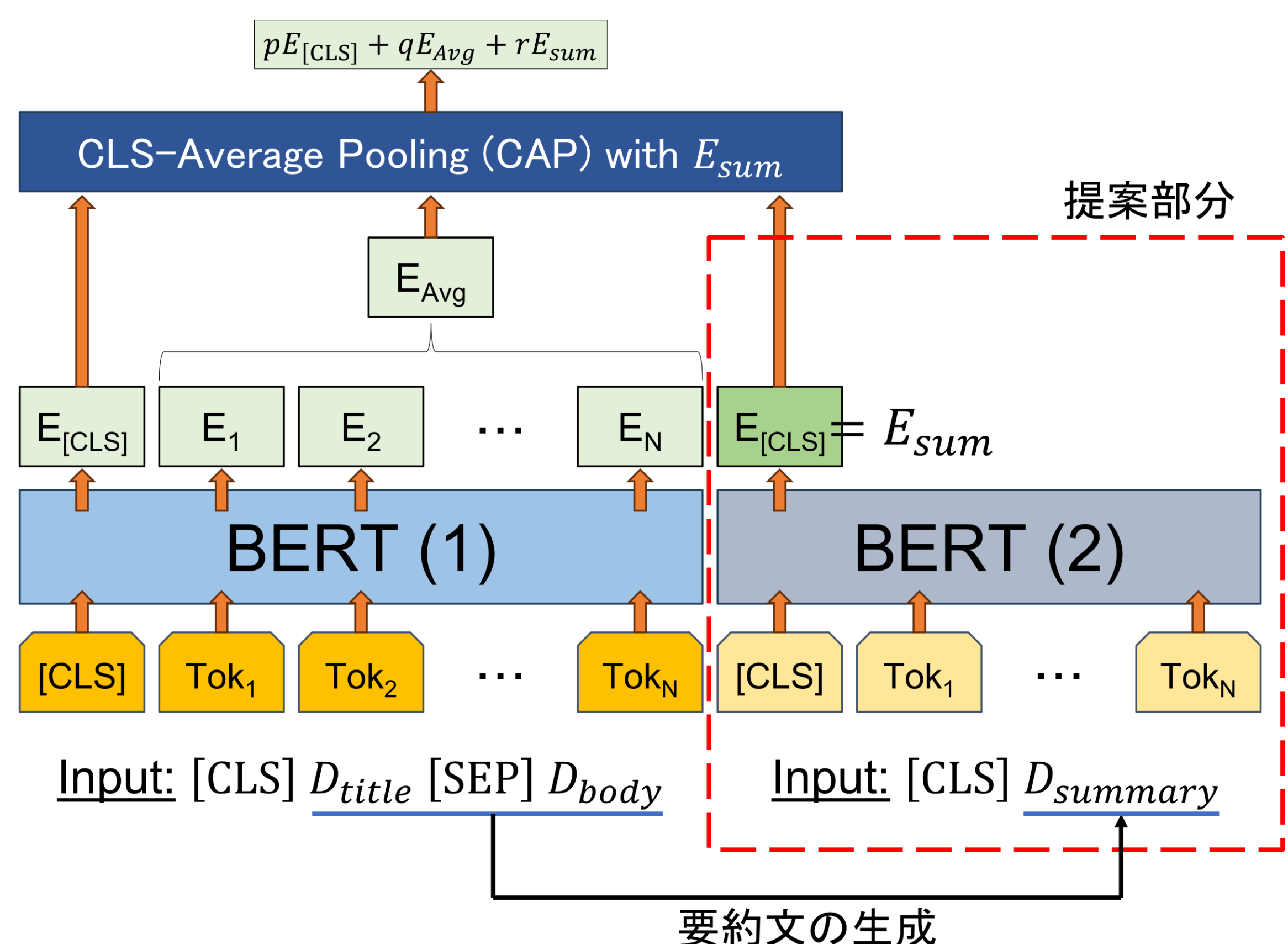
## 提案手法

CAP 層における要約文ベクトル項  $E_{sum}$  の追加

学習可能な和が 1 となるパラメータ  $p, q, r (\geq 0)$   
を用いて  $E_{[CLS]}$ ,  $E_{Avg}$ , 及び  $E_{sum}$  の重み付き和

$$pE_{[CLS]} + qE_{Avg} + rE_{sum}$$

を文の分散表現とする



- 2つの入力系列を独立した訓練済み日本語 BERT モデルへ入力  
日本語 BERT-base モデル (東北大学)
- 各 BERT モデルの最終層の出力より  $E_{[CLS]}$ ,  $E_{Avg}$ , 及び  $E_{sum}$  を  
算出し、提案手法であるプーリング層へ入力

## 実験設定

【livedoor ニュースコーパスデータセットを用いたテキスト分類】

データセットに含まれる各記事データ  $D$  は  
記事タイトル  $D_{title}$ , 記事本文  $D_{body}$ ,  
カテゴリラベル  $D_{label} \in \{0, 1, \dots, 8\}$  を持つ

・要約文  $D_{summary}$  の生成 (PLaMo API<sup>[2]</sup> を利用)

$D_{summary} = \mathcal{L}(D_{prompt})$

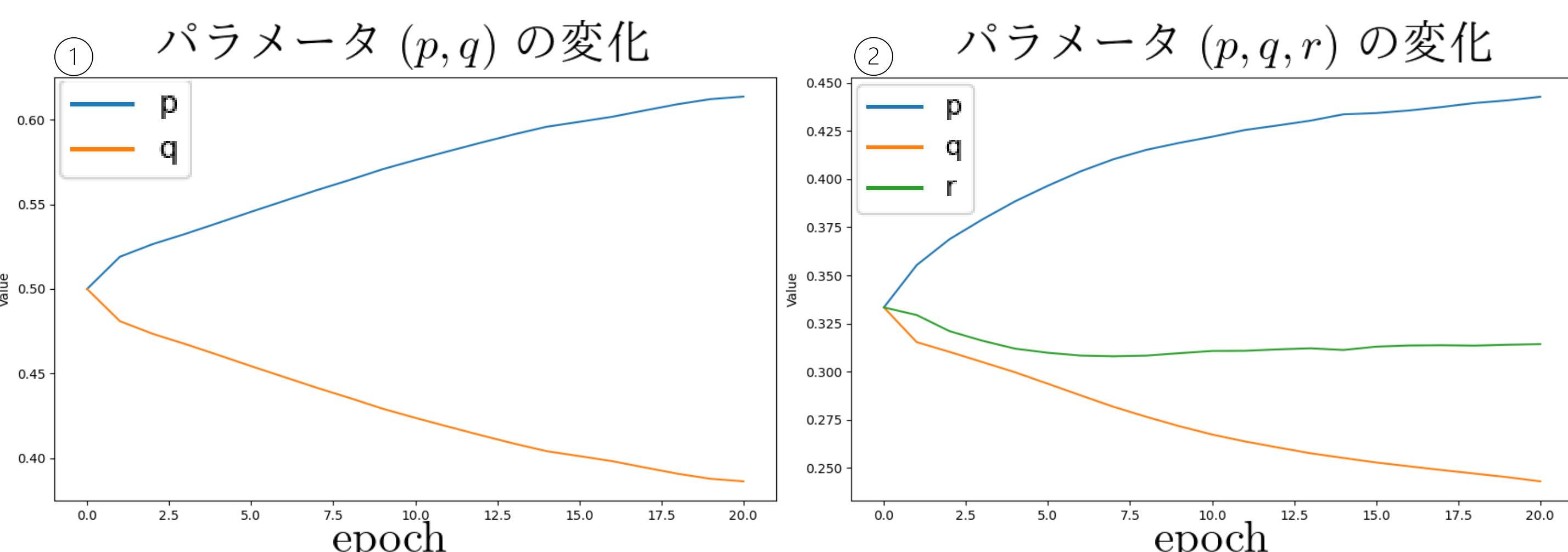
$D_{prompt} = \begin{cases} \text{"role": "system", "content": "あなたは優秀な AI アシスタントです。"} \\ \text{"role": "user", "content": "以下は「 $D_{title}$ 」というタイトルの"} \\ \text{記事の本文です。内容を要約し、} \\ \text{結果のみを出力してください。"} \\ \text{"role": "user", "content": " $D_{body}$ "} \end{cases}$

$\mathcal{L}$ : 大規模言語モデル API (e.g., ChatGPT API, PLaMo API)

学習パラメータ	値		重み初期値
エポック数	20	従来手法	$(p, q) = (\frac{1}{2}, \frac{1}{2})$
バッチサイズ	16	提案手法	$(p, q, r) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$
最適化手法	Adam		
学習率	$3.0 \times 10^{-5}$		
損失関数	Cross Entropy Loss		
分類器	全結合層		

## 実験結果

	Accuracy	F1 (weighted)
① 従来手法 (CAP)	0.9565	0.9560
② 提案手法 (CAP with $E_{sum}$ )	0.9687	0.9683



- Accuracy, F1 値において、  
提案手法の有効性を確認
- $E_{sum}$  による、欠損した原文の  
重要な文脈情報の補完可能性

	トークン数	
	$\geq 512$	$< 512$
Original	3747	2148
Summary	11	5884

## 今後の課題

- 要約文の妥当性, 生成手法改善の検討
- 他のデータセットを用いた提案手法の有効性の確認
- 最適な学習パラメータやアーキテクチャの探索

[1] 大和秀徳. BERT の分散表現に対する学習可能な重みを持つプーリング手法の提案, 大阪公立大学大学院情報科学研究科 基幹情報学専攻 知能情報分野 修士論文, 2024.  
[2] Preferred Elements, <https://plamo.preferredai.jp/>, 2024.