

① 提案手法 (1)

$$C = p E_{cls} + q E_{avg} + t E'_{cls} + s E'_{avg}$$

② 提案手法 (2)

$$C = p E_{cls} + q E_{avg} + t E'_{cls}$$

③ 提案手法 (3)

$$C = p \frac{E_{cls} + E'_{cls}}{2} + q \frac{E_{avg} + E'_{avg}}{2}$$

④ 従来手法 (大森)

(p, s 項をゼロ)

と比較

ランダムに 3回 試行

その平均

$$⑤ C = E'_{cls}$$

$$⑥ C = p E'_{cls} + q E'_{avg}$$

	ACC	F1-noistrel	④との比較 (acc)
①	0.9655 (0.0019)	0.9652 (0.0017)	0.0358 < 0.05
②	0.9655 (0.0023)	0.9652 (0.0022)	0.0416 < 0.05
③	0.9646 (0.0011)	0.9642 (0.0011)	0.05589
④	0.9593 (0.0026)	0.9579 (0.0026)	—
⑤	0.8360 (0.0034)	0.8334 (0.0035)	—
⑥	0.8392 (0.0084)	0.8371 (0.0086)	—

() は標準偏差

400 の CAP ベースに

追加実験

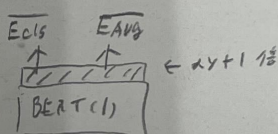
→ 重要語抽出 + 計算量削減を目指す

① 原文と要約文 から tf-idf スコアを算出し

要約文に対する各単語の tf-idf スコアを x_w とする

原文の各単語に対して x_w を辞書参照し、(加減は別) y とする

重みを $\alpha y + 1$ として BERT の最終層出力の x に α 倍する $\alpha(p E_{cls} + q E_{avg})$



	[CLS]	Google	は	スコア	を	作成した	[SEP]	[PAD]
y	0.0	0.4	0.0	0.2	0.0	0.1	0.0	0.0
$\alpha y + 1$ ($\alpha = 5$)	1.0	2.1	1.0	2.0	0.0	1.5	1.0	1.0

この数値を最終層の出力に加算

	Acc	F1	④との比較
$\alpha = 1.0$	0.9601 (0.0026)	0.9598 (0.0026)	0.5183 > 0.05
$\alpha = 5.0$	0.9628 (0.0006)	0.9624 (0.0007)	0.1222 > 0.05
④	0.9583 (0.0026)	0.9579 (0.0026)	—

有意差があるとは

言い換え

仮定に 2 回試行回数で、④ のような性能を得た。

(① ~ ③ と比較するとモデルが 1/2 ほど)

他と比較して

cls より avg を重視して学習がされた、最終的に $p=0.4, q=0.6$

単語間の情報からより正確に取れている

重要語の抽出に貢献