

進捗報告

1 あらすじ

B3 実験において、「自然言語処理と深層学習に基づいた 4 コマ漫画のセリフの感情推定」を行った。

そして、今学期において最初のタスクとしては BERT[1] による分散表現を獲得し、B3 実験と同様に感情推定を行い、結果の比較・考察をすることであった。先週からのタスクとしては BERT の fine-tuning が挙げられる。

2 今週やったこと

- BERT の fine-tuning 実装.
- “1g-hub/bert_practice” レポジトリ作成.

3 BERT の fine-tuning 実装

optuna を用いて最適パラメータを出した後、最適パラメータを用いて fine-tuning 出来るようになった。以下の問題設定で、fine-tuning が動作することを確かめた。実行ソースコードは “1g-hub/takayama/src/ex_2020_4_29.py” にある。

- 入力 は 1 つのセリフ文の単語 id 列, 出力 は感情ラベル (0: '喜楽', 1: 'その他')
- クラス重みは正規化されたラベル逆比
- BERT の事前学習済モデルは日本語 Wikipedia から全 1,800 万文を用いて事前学習させたモデル¹を利用した。
- BERT の出力に 全結合層 (Classification layer) をアダプトし、両方のパラメータを学習。
- optuna でチューニングするのは時短のため learning rate のみ。試行回数は 5 回。
- B3 実験で用いていた learning rate scheduler は不使用。

学習パラメータを表 1, ネットワークの概略図を図 1 に示す。

¹<http://nlp.ist.i.kyoto-u.ac.jp>

表 1: 学習パラメータ

	実験 1・2
epoch	20
batch size	32
loss function	Cross Entropy Loss
optimizer	Adam

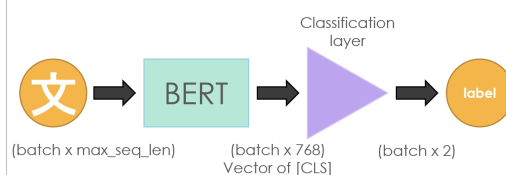


図 1: model

4 BERT fine-tuning テスト動作結果

テスト動作の結果を表 2 に示す。D2V + MLP, D2V + SVM, D2V + RF は B3 実験にて得た、同様の問題設定で分散表現化手法として Doc2Vec を用い、識別器として、3 層からなる多層パーセプトロン (MLP), SVM, RandomForest (RF) を用いた感情推定の結果である。(epoch 数や batch size は異なる.)

計算時間は約 15 sec / epoch であった。あまり良い結果ではなかったが、これがバグによるものなのか、時短方策によるものなのかは定かではない。図 2 にギャグタッチにおける, loss, accuracy, 正例の F1 値の遷移を示している。図 2 を見る限りはうまく学習出来ていない。

5 “1g-hub/bert_practice” レポジトリ作成

金田君と共に BERT の導入, 使い方の指南としてのレポジトリを作成した。随時更新していく。誤っ

た情報などがあれば、指摘をどうかよろしくお願いします。

6 課題 優先度準

- BERT の fine-tuning 推敲. バグ探し.
- 他条件での BERT fine-tuning を行う.
- 新しく作成した d2v モデルと手法を用いた B3 実験の再実験および結果の比較.
- 森先生と大工大の上野先生に 30 話までであるらしい追加データをお願いする.
- Data Augmentation の手法の改善案.

表 2: result

識別器	5 タッチ平均		
	Acc	Recall	F1
D2V + MLP	0.66	0.32	0.31
D2V + SVM	0.65	0.19	0.19
D2V + RF	0.67	0.19	0.20
BERT fine-tuning	0.70	0.30	0.21
ベースライン	0.71	0.00	0.00

参考文献

- [1] Chang M.-W. Lee K. Devlin, J. and K Toutanova.
Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv:1810.04805*, 2018.

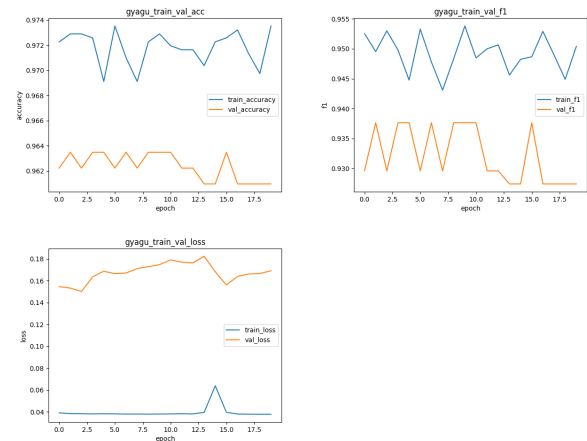


図 2: gyagu graphs