

テキスト含意推定

Textual Entailment

風間 健太郎†

† 東京農工大学工学部 情報工学科

1 背景

テキスト含意推定 (Textual Entailment) は, Argument Mining の中でも近年研究が盛んとなっている分野である. テキスト含意推定では, ある 2 つの文のテキスト間に含意関係があるか矛盾関係があるか, それとも何もないかを推定するタスクである. 具体的には, 「あの人は外科医だ」という文章と, 「あの人は医者だ」という文は後者の文が前者を含意している. 一方で, 「お酒を飲んだ」と「お茶を飲んだ」は矛盾関係を持つ.

テキスト含意推定においては, 2 つの LSTM でエンコードするモデルが考えられてきた. Cocarascu らの研究では, 図 1 のように 2 つの LSTM にそれぞれの文章を入力し, GloVe Embeddings で学習済みの埋め込み表現を用いることで 89% の F 値を達成した.

テキスト含意推定に使われるデータセットとして有名なものとしては SNLI コーパスがある. これは 2 文間に含意関係・矛盾関係・関係なしのアノテーションをしたコーパスである. 一方で, 2018 年に MNLI (Multi-Genre NLI) が登場している. こちらはジャンルごとに違う含意・矛盾関係をアノテーションしたタスクであり, biLSTM でも Accuracy が 67.5% と, より難しいタスクになっている.

テキスト含意推定のためによく使われる手法は以下の通りである.

- CBOW
- BiLSTM
- ESIM

CBOW とは, ある単語の前後の文脈を見て, ある単語を予測するタスクである. BiLSTM とは LSTM の双方向モデルであり, 2 つのテキストを LSTM に同時に入れることで効果を発揮する. ESIM とは 2 つの文関係を見るモデルであり, Attend, Compare, Aggregate の 3 つから成り立つ.

2 問題点

テキスト含意推定における問題点は, 適用できるアーキテクチャが少ないということである. 基本的には, 2 つの LSTM で文章を同時にエンコードし, SoftMax で出力することが一般的である. その他のアーキテクチャについては ESIM 以外まだ考えられていない. テキスト含意推定にも使用できるアーキテクチャとして, (Anirban,

et. al. 2016.) では, 6 つの RNN ベースのモデルを提案している. このモデルを図 2 に示す.

図 2 で注目すべきなのは, 最初の文の LSTM の最後の隠れ層を次の文の埋め込み表現と Concatenate する conditional-input, また最初の文の LSTM の最後の隠れ層を次の LSTM の初期値とする conditional-state-input の 2 つである. 筆者は, Cocarascu らのモデルと conditional-input, conditional-state-input を実装したが, 有意な差は認められなかった.

3 今後の展望

CNN ベースのアーキテクチャを作ることが考えられる. RNN では時系列の処理に限界があるため, アーキテクチャ自身の改善 (それこそ, ICML に通せるような) が今後の展望として見込まれる. そのために, CNN を用いたアーキテクチャが考えられる. 文章の埋め込み表現を畳み込み, 上記の conditional-state-input などを用いることによって, 精度が向上するのではないかと考えられる. また, ESIM についても調査を行い, 理解を深めていく必要がある.

†morio@katfujii.lab.tuat.ac.jp

©藤田桂英研究室. 再配布厳禁.

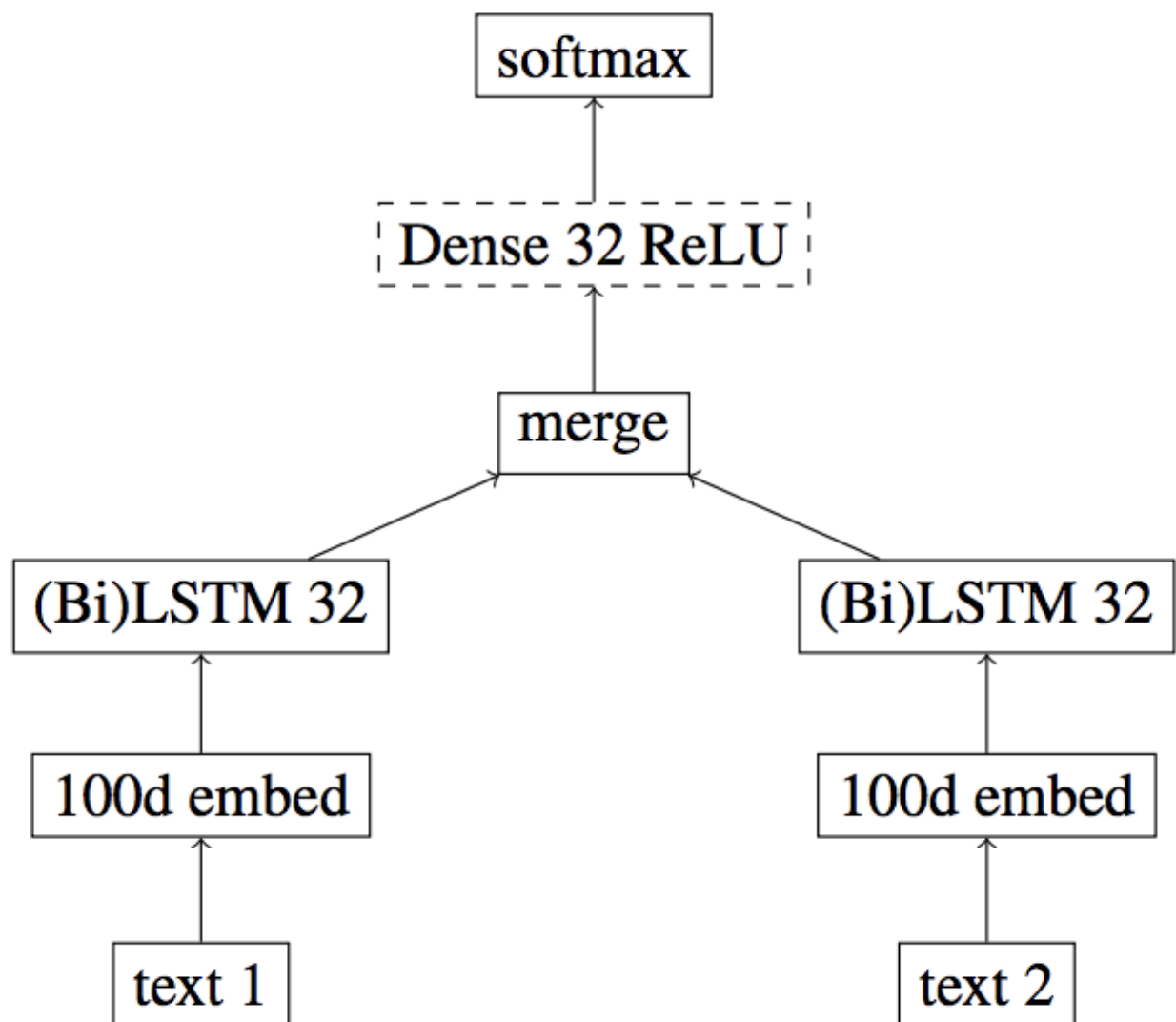
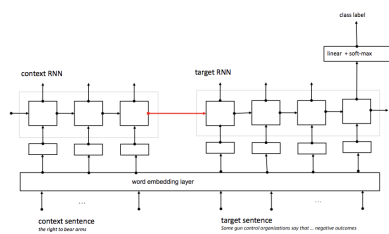
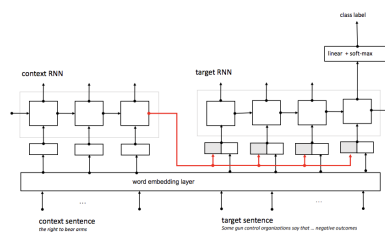


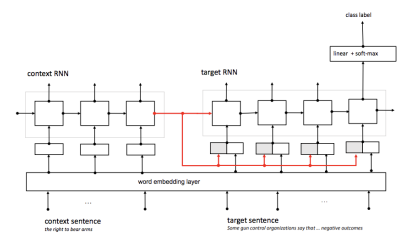
図 1: Cocarascu らのモデル



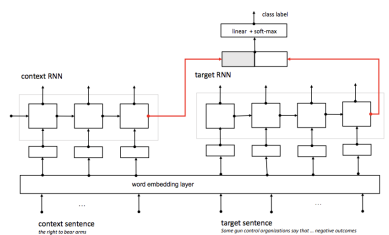
(a) conditional-state



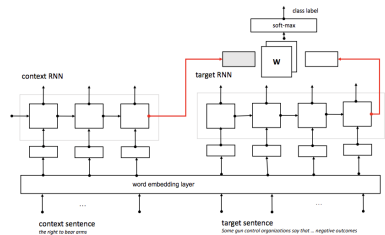
(b) conditional-input



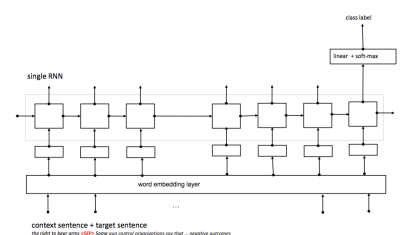
(c) conditional-state-input



(d) concat



(e) bilinear



(f) concat-sentence

図 2: 6 つの RNN モデル