

ニューラルネットワーク言語モデルを用いた 含意文生成システム

柳本研究室 2191104019 趙 曉梅

1. はじめに

含意関係とは、前提となる文 T と文 H が存在した時、文 T が真の場合に文 H が真であると推論できる関係のことである。含意関係を認識することは、自然言語処理における情報検索、質問応答、情報抽出などの様々なタスクにおいて必要とされる技術である。Dagan ら[1]は、推論ルールを用いて含意関係の抽出を行なっている。

しかし、含意関係にある 2 文には様々なものがあり、ほぼ前提文と同じものであっても含意文と判断されてしまう。例えば、以下の例を用いると、含意文 1 は単語を前提文から単語を削除して作成されているが、含意文 2 は異なった語彙で作成されている。

前提文 He went to the library again to borrow books.
含意文 1 He went to borrow books.(類似度 0.8398)
含意文 2 This is not his first visit to the library. (類似度 0.7145)

本論文では、できるだけ前提文と語彙が重ならないような含意文を作成することを目的とする。含意生成には、機械翻訳でよく用いられる Attention 付きの Sequence-to-Sequence モデルを用いることとする。さらに、前提文と含意文との語彙の重複を制限するため、前提文と生成された含意文の類似度に基づいた損失を定義し、それを用いることで目的を達成する。提案手法を評価するために Stanford Natural Language Inference(SNLI)コーパスを用いて実験を行い、前提文と語彙を共有するような含意文を抑えることができた。

2. 関連研究

Attention based Sequence-to-Sequence モデル[2][3]は幅広い分野の自然言語処理タスクに応用され、含意文生成タスクのようなテキスト生成タスクにおいて高い性能を示しているモデルである。

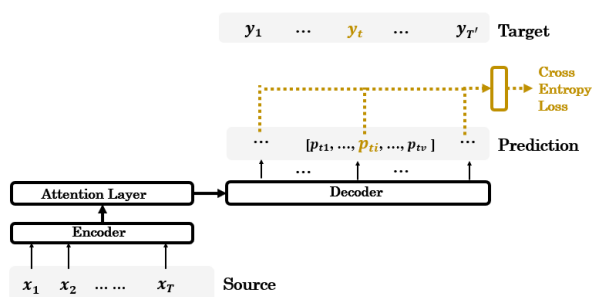


図 1: ベースライン手法の構成

Attention based Sequence-to-Sequence モデルでは、前提文となる入力 Encoder で圧縮され、Decoder に渡される。Decoder では、Encoder からの情報と前の時刻の出力を計算して含意文を予測する。コーパスの含意文を教師信号として予測された含意文に交差エントロピー誤差を用いて学習される。この学習により、コーパスとして用意された含意文に生成される含意文の質は依存することとなる。

3. 提案手法

3.1. 類似損失 (Similarity Loss)

前提文と生成された含意文との語彙の重なりを小さくすることで、前提文と異なる含意文の生成を目指す。このため、入力文と出力文の類似度を損失に組み込むこととする。図 2 に提案システムの構成を示す。

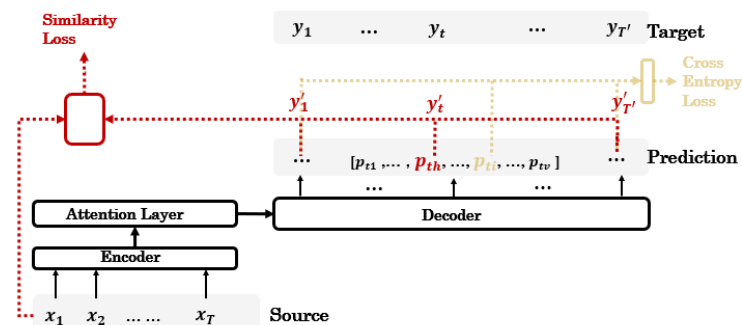


図 2: 提案手法の構成

予測した含意文 $[y'_1, \dots, y'_{T'}]$ と入力である前提文 $[x_1, \dots, x_T]$ の類似損失を計算する。そして、類似損失と交差エントロピー誤差損失の和を訓練データにおける誤差関数として用いる。

類似損失は $[y'_1, \dots, y'_{T'}]$ と $[x_1, \dots, x_T]$ をもとに計算する。 $[y'_1, \dots, y'_{T'}]$ と $[x_1, \dots, x_T]$ で共有されているトークン数を計算して、 $[y'_1, \dots, y'_{T'}]$ の長さ T' で割ることで、類似損失を定義する。例えば、図 3 では、含意文の長さは 9 で、前提文と共有されているトークンは 3 つであるため、含意文の長さで割ると、類似損失は $\frac{1}{3}$ となる。

前提文 He went to the library again to borrow books

共有トークン 3 つ

含意文 This is not his first visit to the library

長さ 9

図 3: 類似損失の例

類似損失と交差エントロピー誤差損失の和を訓練データにおける最終的な誤差関数とする。

$$\text{Total Loss} = \text{Similarity Loss}([x_1, \dots, x_T], [y'_1, \dots, y'_{T'}])$$

$$+ \text{Cross Entropy Loss}$$

(1)

こうすることで、入力文と近い含意文が生成された場合には誤差が大きくなり、出来るだけ入力文と異なった含意文が生成されるようになる。

3.2. 評価手法

生成した含意文と入力となる前提文の文間類似度を求めるために、訓練済みの BERT[4]を用いて、前提文と生成された含意文をベクトルで表現し、その類似度をコサイン類似度により計算する。単語埋め込みサイズは 768 である。生成した含意文と入力となる前提文をそれぞれ BERT に入力して、最後の 4 層の出力の平均ベクトルを文ベクトルにする。生成し

た含意文と入力となる前提文の文ベクトルをもとに、コサイン類似度を計算する。これを類似度についての評価指標にする。

4. 実験

4.1. データセット

Stanford Natural Language Inference(SNLI)コーパスは Samuel R ら[5]によって作成されたコーパスである。自然言語推論タスクでよく使用されており、ラベル付きのコーパスとして代表的なものである。2 文の関係は、「含意(entailment)」「中立(neutral)」「矛盾(contradiction)」という 3 つのラベルが付与されている。実験では、含意文生成をタスクにするため、「含意」ラベルが付与された含意関係である前提と仮説のペアのみ取り出して、実験に用いた。取り出したデータ数は 190,113 であり、訓練データが 183,416 件、検証データが 3,328 件、テストデータが 3,360 件に分割されていた。

類似度問題の原因の一つは、データセットにあると考えられる。SNLIにおいて、訓練データとなる前提文と含意文の文間類似度は平均で 0.7696 である。データセットにおける類似度は[0.2, 1]の範囲に分布され、類似度が[0.5, 1]範囲内のデータは全体の 99.446%を占めている。

4.2. 実験設定

エンコーダとデコーダの隠れ層はそれぞれ 2 層の Bi-GRU と 2 層の GRU からなるものであり、隠れ層のサイズと単語埋め込みのサイズは 1,024 に設定し、0.2 の Dropout を採用した。注意層では、「Global Attention」と「General Attention Score」を採用した。訓練の段階で、オプティマイザは「Adam」を使用した。バッチサイズは 32 で、GPU 上で学習を行なった。

表 1: 実験設定

アイテム	ベースライン手法	提案手法
Word Embedding	1024	1024
Encoder	2 層 Bi-GRU	2 層 Bi-GRU
Decoder	2 層 GRU	2 層 GRU
GRU Hidden Size	1024	1024
Dropout	0.2	0.2
Learning Rate	0.0001	0.0001
Optimizer	Adam	Adam
Regularization	L2	L2
Batch Size	32	32
Loss Function	Cross Entropy Loss	Cross Entropy Loss + Similarity Loss

5. 結果

提案手法とベースライン手法をそれぞれ訓練データで 10 回学習させて、テストデータによる含意文の生成を行い、前提文と含意文の類似度の平均を表 2 に示す。平均類似度の点からはベースラインと提案手法に大きな差は現れなかった。

表 2: 訓練結果

	ベースライン	提案手法
平均類似度	0.7517	0.7580

両手法で生成された含意文と前提文の類似度の分布を図 3 と図 4 に示す。類似度が 0.9 以上のものは減少していること

がわかる。

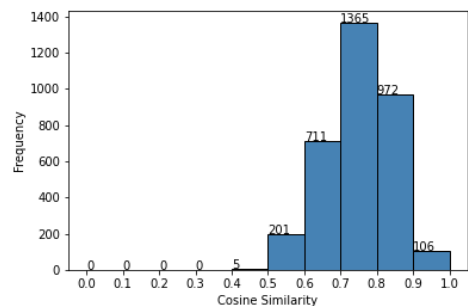


図 3: ベースライン手法によるテストデータの類似度分布

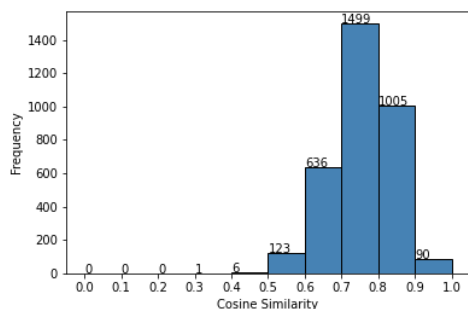


図 4: 提案手法によるテストデータの類似度分布

6. おわりに

本論文では、前提文と生成された含意文の語彙が重ならないように、類似損失を導入した手法を提案した。これにより、前提分の語彙をそのまま使ったような含意文の生成を抑えることができた。

今後の課題としては、全体的に前提文と含意文との語彙の重なりを抑えるようにすることを目指す。また、学習回数を増やすことでどのように変化するかについても検討を行う。

参考文献

- [1] Dagan I, Glickman O. Probabilistic textual entailment: Generic applied modeling of language variability. In *Proceedings of the PASCAL Workshop on Learning Methods for Text Understanding and Mining, Grenoble*, pp26-29, 2004
- [2] I. Sutskever, O. Vinyals, QV Le. Sequence to Sequence Learning with Neural Networks. In *Proceeding of NIPS*, pp.3104- 3112,2014.
- [3] M. Loung, H. Pham, C. D. Manning. Effective Approaches to Attention-based Neural Machine Translation. In *Proceeding of EMNL*, pp.1412- 1421,2015.
- [4] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceeding of Association for Computational Linguistics*, pp.4171-4186, 2019.
- [5] Samuel R. Bowman, Gabor Angeli, Christopher Potts, Christopher D. Manning. A large annotated corpus for learning natural language inference. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, Lisbon, Portugal, pp.632-642, Sep.2015.