

自然言語処理における連続的統語構造変換と対照学習の拡張

理学専攻・情報科学コース 2040639 折口 希実

1 はじめに

構造情報は有益な情報である一方、扱いつらさから学習においてあまり考慮されていない場合が多くある。深層学習のフレームワークで構造情報を処理する方法として、Kusnerら [1] は離散な構造を持つデータを連続的な値、つまり潜在空間における埋め込みベクトルとして扱うことができる、変分自己符号化器 (VAE) モデル [2] に基づいた新しい VAE である Grammar Variational Autoencoder (GVAE) を提案している。本研究では、GVAE モデルを改良し、自然言語文の統語構造を連続的に変換可能にし、ガウス分布によって表された潜在空間からサンプリングする点の変化によって連続的に統語構造を変換可能にする手法を開発する。さらに、潜在空間にエンコードされたベクトルを使用し、統語構造の観点から文の類似性を測定する方法を提案する。類似性の測定手段として同じ画像同士の特徴量を最大化し異なる画像同士の特徴量を最小化する対照学習はとて有効であり、学習テクニックを大幅に進歩させ優れたパフォーマンスを実現した。このフレームワークを自然言語へと導入した SimCSE を用い、さらなる性能の向上へ向けて敵対的なサンプルを用いた対照学習を行った。負例の生成方法についても着目し、さまざまな工夫を行いながら高品質な文の埋め込み表現の学習を可能とする学習の実現を目指す。

2 文法自己変分符号化器

Kusner ら [1] が提案した自己符号化器 (GVAE) は、入力として離散なデータを扱うことができる VAE モデルの一種であり、文法の生成規則も使用することが可能である。潜在空間のサンプリング点を僅かに変えることで離散情報である自然言語の統語構造をシームレスに変換することができる。彼らはその適用例として、分子の化学構造を ASCII 符号の英数字で文字列化した表記法である SMILES 記法を用いた分子構造に対して連続的な表現および変換方法を示している。

統語構造に対する GVAE GVAE を適用して自然言語の統語構造の類似性を潜在空間において計測可能としつつ、潜在空間でのサンプリングする値の変化を通じて与えられた文の統語構造の類似関係を制御し統語構造を変換する。GVAE の処理の流れは図 1 の通りである。詳細なプロセスについては本稿では省略する。

GVAE のフレームワークを自然言語の統語構造に適用する場合、Kusner ら [1] で示される SMILES 記法に基づいた分子に適用する場合とは異なり、自然言語文の統語構造は非終端記号の数と文法の生成規則数の両方の点でより複雑となる。解析結果は複数存在するため、より確からしい結果を選択する必要がある。本研究では文法規則として確率文脈自由文法 (Probabilistic Context-free Grammar ; PCFG) を採用し、viterbi parser¹を用いて構文解析を行う。また、ガウス分布となっている潜在空間のサンプリング点を

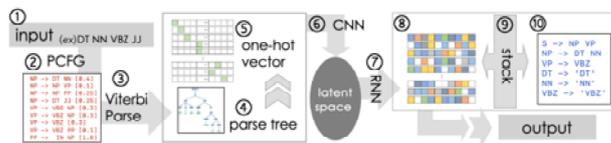


図 1: 自然言語の統語構造に対する GVAE のプロセス

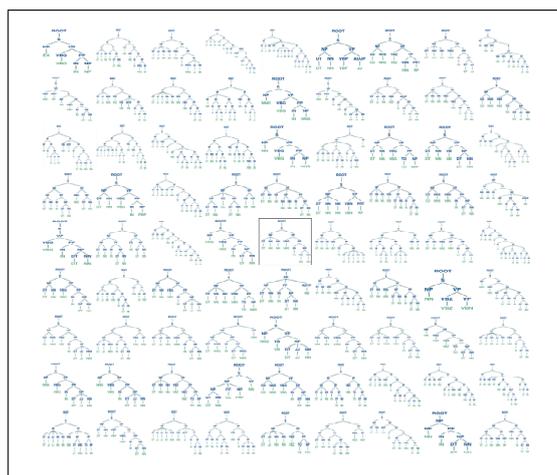


図 3: 統語構造 1 の連続的な変換 (サンプリング点は平均から 0.01 刻みで移動)

分散の変化によってシームレスな生成を可能とする。

2.1 実験

本研究では、潜在空間で表されるガウス分布からのサンプリング点を徐々に変化させることによる統語構造の連続的変換および異なる統語構造の類似性を測定する方法を検証するための実験を行なった。データセットとして、Microsoft COCO データセット²を用い、10 語程度の平易な文である 6,668 の文を選択し、

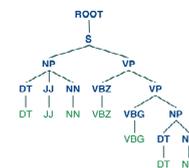


図 2: 統語構造 1

StanfordCoreNLP³を用いて文を解析を行い、GVAE で使用される生成規則を構築した。また、PCFG を構築するために生成規則とその確率を計算して導出し、解析に必要な 320 の生成規則を取得した。統語構造 1 'DT JJ NN VBZ VBG DT NN' (図 2 参照) を GVAE の入力とし、図 3 は、その出力結果を表している。評価指標がなく、視覚的な評価となってしまうが滑らかな変換も見受けられる一方、突飛で不完全な変換となってしまう部分が存在することもわかる。

統語構造の類似性を、コサイン類似度を用いて潜在変数 z で測定した。表 2 は、統語構造を僅かに変化させていった際のコサイン類似度を表している。例文を与え、統語構造の変化とコサイン類似度の変化の関係について調べた。

¹https://www.nltk.org/_modules/nltk/parse/viterbi.html

²<https://cocodataset.org/home>

³<https://stanfordnlp.github.io/CoreNLP/>

表 1: 実験結果 (STS タスク)

Model	STS12	STS13	STS14	STS15	STS16	STS-B	SICK-R	Avg.
SimCSE	66.07	79.99	71.14	78.91	77.74	75.16	69.39	74.06
+GS-InfoNCE(GS)	67.24	81.41	73.05	79.21	78.38	75.30	69.59	74.88
ours [+adversarial examples(ae)]	68.88	82.15	73.71	80.97	79.13	77.98	69.71	76.08
ours [+GS+ae]	66.78	79.49	72.00	79.46	78.80	76.02	70.28	74.69

表 2: 統語構造の違いによる cos 類似度の変化

Sentence examples	統語構造	cos 類似度
A boy is riding a horse.	DT NN VBZ VBG DT NN	—
A girl is riding an elephant.	DT NN VBZ VBG DT NN	1
Two boys are riding a horse.	CD NNS VBZ VBG DT NN	0.99999
A boy rides a horse.	DT NN VBZ DT NN	0.99992
A young boy is riding a horse.	DT JJ NN VBZ VBG DT NN	0.99986

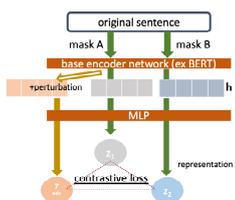


図 4: 提案モデル

3 対照学習の性能向上に向けて

3.1 自然言語処理と対照学習

対照学習とは機械学習の手法の一つであり、データ同士の比較によって大量のデータをラベル付けることなく学習することが可能である。対照学習では基準の軸となるデータであるアンカーとそれに似たデータである正例を近くに、異なるデータである負例は遠くに埋め込まれるように学習を行う。

SimCSE は [3] によって提案された埋め込み学習のフレームワークである。データの拡張にドロップアウトマスクのみを用いたシンプルな仕組みだが、SimCSE モデルは既存手法の精度を更新する十分な精度を実現している。同じ文を BERT 等の自然言語モデルに 2 度渡してベクトルとし、得られた 2 つのベクトルを正例のペアとする。このとき 2 つのベクトルは同じ文から取得したものであるが、ドロップアウトマスクを用いることで僅かに異なった表現となっている。通常 BERT への入力の際、複数の文を同時にベクトル化しているが、そのときのミニバッチ内の他の文については負例として扱うこととする。損失関数である NT-Xent loss を用いて正例のペアのソフトマックスを 1 に近づけるように、そうでないものを 0 に近づけるように学習することで対照学習を行うことができる。

3.2 提案手法

敵対的サンプル 敵対的サンプルは、機械学習モデルに誤った予測をさせてしまう、微小な摂動を加えたサンプルのことである。作成方法として代表的なもの一つに Fast Gradient Sign Method (FGSM) がある [4]。モデルを誤認識させてしまう敵対的サンプルを混ぜて学習を行うことでより堅牢なモデルを得られることが報告されている。

本研究においては敵対的サンプルを用いて対照学習を行うモデルを構築した (図 4 参照)。

4 実験

提案モデルを含め実験を行い、対照学習の性能を向上させる方法を検証する。評価方法についてだが、先行研究と同様に SentEval: evaluation toolkit for sentence embeddings⁴を用いた。SentEval の 7 つの STS (Semantic Textual Similarity) のタスク⁵によって文の埋め込みを評価することができる。文章の等価性を評価するタスクであり、2 つの文の関係について類似度を取り評価する。対照学習は「似た表現を近く同士に、異なる表現を遠くへ」と学習させるものであるため、類似を評価する STS タスクを用いることで文の埋め込み表現のパフォーマンスについて評価が可能である。

パラメータは [3] を参考にして提案モデルで対照学習を行い、それを用いて実験をした STS タスクの結果を先行研究らの追実験の結果と比較し表 1 に示す。これらの結果は計算コストの軽減の為に BERT ベースの学習済み言語モデルを用いて行っている。全体的に先行研究の結果を上回り、敵対的サンプルの導入が対照学習において有益であることがわかった。

また、敵対的サンプルの導入だけでなく負例の生成方法を工夫したモデルや損失関数の変更をしてモデルを構築し実験等を行なった。ここでは要旨のため割愛するが論文にて結果とともに記す。

5 おわりに

本稿では行った 2 つの研究について記している。2 節では、自然言語の統語構造を対象とした文法変分自己符号化器を作成し、構造の滑らかな変換を行った。類似性などを測る指標となり得る可能性が感じられた。3 節では、対照学習の性能向上に向けて敵対的サンプルを用いるなど様々な工夫をして実験を行った。下流タスクにおいて先行研究を上回る結果を出すことができ、提案手法が有効であることを示した。今後の課題としては、より性能を向上することができる堅牢なモデルを構築したい。

参考文献

- [1] Matt J. Kusner, Brooks Paige, and José Miguel Hernández-Lobato. Grammar variational autoencoder. Vol. 70 of *Proceedings of Machine Learning Research*, pp. 1945–1954, International Convention Centre, Sydney, Australia, 06–11 Aug 2017. PMLR.
- [2] Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. In Yoshua Bengio and Yann LeCun, editors, *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14–16, 2014, Conference Track Proceedings*, 2014.
- [3] Tianyu Gao, Xingcheng Yao, and Danqi Chen. SimCSE: Simple contrastive learning of sentence embeddings. In *Empirical Methods in Natural Language Processing (EMNLP)*, 2021.
- [4] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples, 2015.

⁴<https://github.com/facebookresearch/SentEval>

⁵STS12, STS13, STS14, STS15, STS16, STS Benchmark, SICK Relations