

# 文対モデリングのための言い換えに基づく 対照学習に関する研究

---

愛媛大学工学部工学科コンピュータ科学コース

人工知能研究室 学部4年 杉山 誠治

(指導教員：二宮 崇, 梶原 智之)

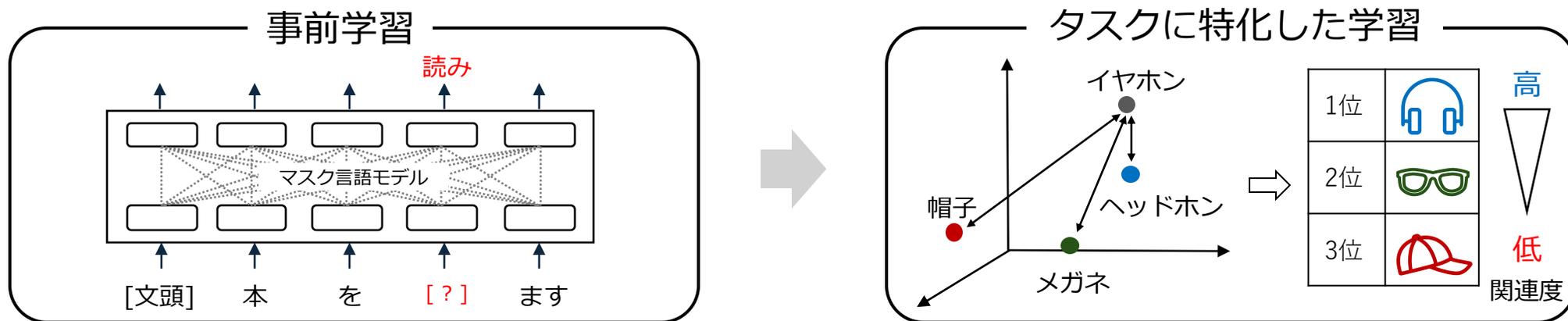
卒業論文発表会

# 背景：マスク言語モデルを利用した学習

- マスク言語モデル<sup>[1,2]</sup>によって、文や単語をベクトルに変換できる



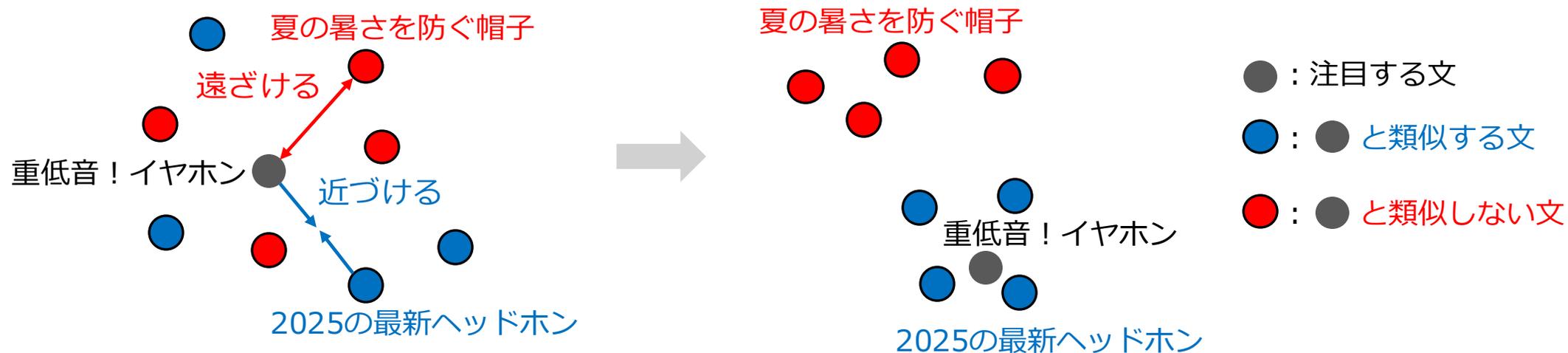
- 事前学習をした後、タスクに特化した学習をしてマスク言語モデルを利用



# 課題：類似性の低い文のベクトル

事前学習では、必ずしも類似する文のベクトルを近づけていない[3]

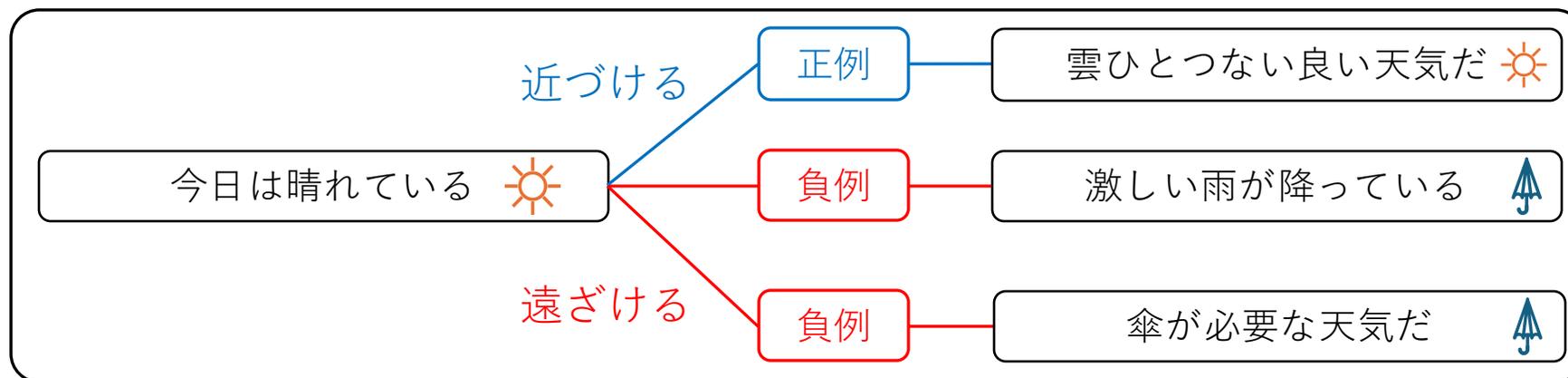
- 類似する文のベクトルを近づけることで、2文間の類似性を捉えられる
- 文対の関連度を扱うタスクの性能を改善できる



# 関連研究：類似する文のベクトルを近づける

類似する文ベクトルを近づけ、類似しない文ベクトルを遠ざける学習が有効

- **対照学習** [4-6] と言われ、人手で作成されたデータセットを使用
- 意味的に等しい文ペアを**正例**、意味的に等しくない文ペアを**負例**として学習



しかし、英語以外の言語において大規模なデータセットは存在しない

# 本研究の概要

1. 言い換え辞書によって、大規模な類似する文（言い換え文）を自動作成  
様々な言語の言い換え辞書<sup>[7]</sup>が存在し、多言語に応用できる手法



2. 事前学習の後に、言い換え文のベクトルを近づける対照学習を追加  
言語モデルが 言い換え元の文 と 言い換え文 の2文間の類似性を捉えられる



# 1-1 : 言い換え辞書によるデータ構築

言い換え辞書を使って言い換え文を作成

- 言い換え文は正例ペアの候補文となる
- 2単語以上のフレーズも考慮

私はこの本の**執筆者**だ。

言い換え元	言い換え先	言い換え確率
<b>執筆者</b>	<b>作家</b>	0.07
<b>執筆者</b>	<b>著者</b>	0.39
<b>執筆者</b>	<b>筆頭 著者</b>	0.01
<b>執筆者</b>	<b>著者は</b>	0.02

候補文を作成



正例ペアの候補文

私はこの本の**作家**だ。  
私はこの本の**著者**だ。  
私はこの本の**筆頭 著者**だ。  
私はこの本の**著者は**だ。

# 1-1 : 言い換え辞書によるデータ構築

言い換え辞書を使って言い換え文を作成

- 言い換え文は正例ペアの候補文となる
- 2単語以上のフレーズも考慮

私はこの本の **執筆者** だ。

言い換え元	言い換え先	言い換え確率		正例ペアの候補文
執筆者	作家	0.07	候補文を作成 →	私はこの本の <b>作家</b> だ。
執筆者	著者	0.39		私はこの本の <b>著者</b> だ。
執筆者	筆頭 著者	0.01		私はこの本の <b>筆頭 著者</b> だ。
執筆者	著者は	0.02		私はこの本の <b>著者は</b> だ。

- ✓ 文法的に**不適切な文**
- ✓ 学習時に**ノイズ**となってしまう

# 1-2：最も流暢な言い換え文の選択

全ての候補文に対してperplexityを計算

- 文の流暢性を推定
- GPT-2<sup>[8]</sup>のニューラル言語モデルを使用
- 言い換え文として最も流暢な1文を選択

$$\text{perplexity}(x) = \exp\left(-\frac{1}{|x|} \sum_{i=1}^{|x|} \log P(x_i|x_{<i})\right)$$

$i$  番目より前の単語が与えられた際に、 $i$  番目の単語を出力する条件付き確率

正例ペアの候補文	perplexity ↓
私はこの本の <b>作家</b> だ。	317.8
私はこの本の <b>著者</b> だ。	156.3
私はこの本の <b>筆頭 著者</b> だ。	143.4
私はこの本の <b>著者</b> はだ。	333.7

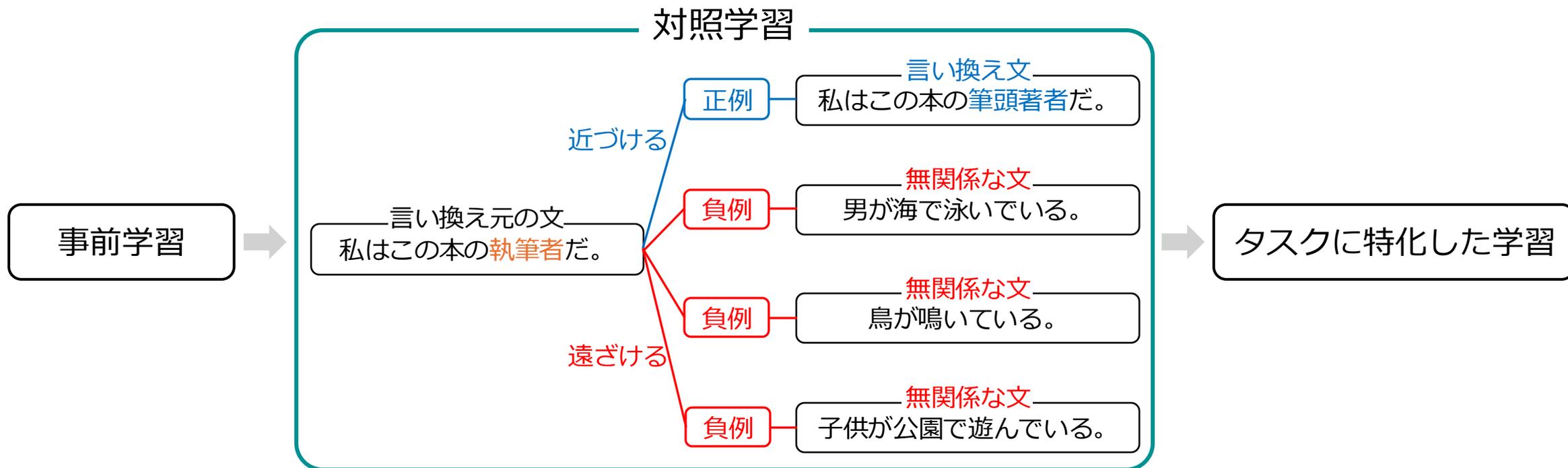
文の流暢性を測定



## 2 : 言い換え文を基に対照学習

事前学習の後に、言い換え文のベクトルを近づける対照学習を追加

- 言い換え文を**正例**、データセット内の無関係な文を**負例**とする
- 言い換え文のベクトルを**近づけ**、無関係な文のベクトルを**遠ざける**



# 実験設定：文対間の類似性を捉えるタスク

- データセット：2文をマスク言語モデルに入力し、その類似性を捉えるタスク

タスク	日本語	英語	評価指標
商品検索	Shopping Queries <sup>[9]</sup>	Shopping Queries	Micro-F1
類似度推定	JSTS <sup>[10]</sup> , JSICK <sup>[11]</sup>	STS-B <sup>[13]</sup> , SICK <sup>[14]</sup>	Spearman
含意関係認識	JNLI <sup>[10]</sup> , JSICK	SNLI <sup>[15]</sup> , SICK	Macro-F1
言い換え認識	PAWS-X <sup>[12]</sup>	PAWS <sup>[16]</sup>	Macro-F1

- マスク言語モデル：日本語 RoBERTa<sup>[17]</sup> / 英語 BERT<sup>[18]</sup>
- 比較手法
  - 教師なしSimCSE：同一文を正例として対照学習
  - 教師ありSimCSE：人手で作成されたデータセットを使用して対照学習

# 実験結果（日本語）

- 言い換え認識以外の3タスクで追加学習なしベースライン以上の性能を達成
- 全タスクの平均性能で、提案手法が最高性能を達成

タスク	商品検索	類似度推定		含意関係認識		言い換え認識	
日本語	Shopping Queries	JSTS	JSICK	JNLI	JSICK	PAWS-X	平均
追加学習なし	0.576	0.859	0.890	0.785	0.839	0.793	0.790
教師なしSimCSE	<b>0.587</b>	<b>0.861</b>	0.886	0.781	0.837	0.790	0.790
教師ありSimCSE	0.576	0.825	0.886	<b>0.843</b>	0.843	<b>0.800</b>	0.796
提案手法	<b>0.587</b>	<b>0.861</b>	<b>0.896</b>	0.828	<b>0.856</b>	0.791	<b>0.803</b>

# 実験結果（英語）

- 全タスクで追加学習なしベースライン以上の性能を達成
- 全タスクの平均性能で、提案手法が最高性能を達成

タスク	商品検索	類似度推定		含意関係認識		言い換え認識	
英語	Shopping Queries	STS-B	SICK	SNLI	SICK	PAWS	平均
追加学習なし	0.654	0.824	0.815	<b>0.904</b>	0.858	0.913	0.828
教師なしSimCSE	<b>0.655</b>	0.830	0.806	<b>0.904</b>	<b>0.868</b>	<b>0.918</b>	0.830
教師ありSimCSE	<b>0.655</b>	<b>0.857</b>	0.824	0.901	0.865	0.913	0.836
提案手法	<b>0.655</b>	0.841	<b>0.842</b>	<b>0.904</b>	0.866	<b>0.918</b>	<b>0.838</b>

# まとめ：文対モデリングのための言い換えに基づく対照学習に関する研究

- 背景：事前学習後に、タスクに特化した学習をしてマスク言語モデルを利用
- 課題：事前学習では必ずしも類似する文のベクトルを近づけていない
- 手法：言い換え文を用いて対照学習して2文間の類似性を捉える
- 結果：日英の両言語において4タスクの平均性能で提案手法が最高性能を達成

## 対外発表

- 杉山誠治, 近藤里咲, 梶原智之, 二宮崇.  
商品検索のための追加事前学習としての言い換えに基づく対照学習.  
NLP若手の会第19回シンポジウム, S1-P27, 2024年9月.
- 杉山誠治, 近藤里咲, 梶原智之, 二宮崇.  
文対モデリングのための言い換えに基づく対照学習.  
言語処理学会第31回年次大会, pp.3238-3242, 2025年3月. (発表予定)
- 国際会議に投稿し、査読中

# 参考文献

---

- [1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pretraining of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 4171-4186, 2019.
- [2] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv:1907.11692, 2019.
- [3] Bohan Li, Hao Zhou, Junxian He, Mingxuan Wang, Yiming Yang, and Lei Li. On the Sentence Embeddings from Pre-trained Language Models. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pp. 9119-9130, 2020.
- [4] Tianyu Gao, Xingcheng Yao, and Danqi Chen. SimCSE: Simple Contrastive Learning of Sentence Embeddings. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, pp. 6894-6910, 2021.
- [5] Yung-Sung Chuang, Rumen Dangovski, Hongyin Luo, Yang Zhang, Shiyu Chang, Marin Soljagic, Shang-Wen Li, Scott Yih, Yoon Kim, and James Glass. DiffCSE: Difference-based Contrastive Learning for Sentence Embeddings. In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 4207-4218, 2022.
- [6] Jiduan Liu, Jiahao Liu, Qifan Wang, Jingang Wang, Wei Wu, Yunsen Xian, Dongyan Zhao, Kai Chen, and Rui Yan. RankCSE: Unsupervised Sentence Representations Learning via Learning to Rank. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, pp. 13785-13802, 2023.
- [7] Juri Ganitkevitch and Chris Callison-Burch. The Multilingual Paraphrase Database. In Proceedings of the Ninth International Conference on Language Resources and Evaluation, pp. 4276-4283, 2014.
- [8] Alec Radford, Jeff Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language Models Are Unsupervised Multitask Learners. 2019.
- [9] Chandan K. Reddy, Lluís Màrquez, Fran Valero, Nikhil Rao, Hugo Zaragoza, Sambaran Bandyopadhyay, Arnab Biswas, Anlu Xing, and Karthik Subbian. Shopping Queries Dataset: A Large-Scale ESCI Benchmark for Improving Product Search. arXiv:2206.06588, 2022.

# 参考文献

---

- [10] Kentaro Kurihara, Daisuke Kawahara, and Tomohide Shibata. JGLUE: Japanese General Language Understanding Evaluation. In Proceedings of the Thirteenth Language Resources and Evaluation Conference, pp. 2957-2966, 2022.
- [11] Hitomi Yanaka and Koji Mineshima. Compositional Evaluation on Japanese Textual Entailment and Similarity. Transactions of the Association for Computational Linguistics, Vol. 10, pp. 1266-1284, 2022.
- [12] Yinfei Yang, Yuan Zhang, Chris Tar, and Jason Baldridge. PAWS-X: A Cross-lingual Adversarial Dataset for Paraphrase Identification. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, pp. 3687-3692, 2019.
- [13] Daniel Cer, Mona Diab, Eneko Agirre, Iñigo Lopez-Gazpio, and Lucia Specia. SemEval-2017 Task 1: Semantic Textual Similarity Multilingual and Crosslingual Focused Evaluation. In Proceedings of the 11th International Workshop on Semantic Evaluation, pp. 1-14, 2017.
- [14] Marco Marelli, Stefano Menini, Marco Baroni, Luisa Bentivogli, Raffaella Bernardi, and Roberto Zamparelli. A SICK Cure for the Evaluation of Compositional Distributional Semantic Models. In Proceedings of the Ninth International Conference on Language Resources and Evaluation, pp. 216-223, 2014.
- [15] Samuel R. Bowman, Gabor Angeli, Christopher Potts, and 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, pp. 632-642, 2015.
- [16] Yuan Zhang, Jason Baldridge, and Luheng He. PAWS: Paraphrase Adversaries from Word Scrambling. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 1298-1308, 2019.
- [17] <https://huggingface.co/rinna/japanese-roberta-base>
- [18] <https://huggingface.co/google-bert/bert-base-uncased>