

市民協働支援のための BERT を用いた Web 記事への社会問題タグ付け手法の試作

長谷川 徳賢[†] 白松 俊[‡]

名古屋工業大学 大学院工学研究科[†]

1 はじめに

近年, COVID-19, 高齢化の様な持続可能性を脅かす様々な社会問題が増加している. これらの社会問題に対応するためには市民と行政が一体となって社会問題に取り組む必要があると考える. 社会問題にどのように取り組むかを議論するためには, 既存の活動を調査することが重要である. 本研究では Web 上の社会問題記事に対して探索的閲覧が可能なインターフェイスの開発のために, Wikidata と Wikipedia を用いて社会問題コーパスを作成し, 汎用言語表現モデル BERT [1] を用いて Web 記事に社会問題タグを付与する手法を提案する.

2 社会問題のコーパスの作成

Web 記事中に含まれる文ごとに, その文が社会問題に言及しているかを判定し, 言及している場合にはその社会問題の種別を自動判定することで, 当該文への自動タグ付けを行う. そのための訓練コーパスを作成する. 具体的には, Wikidata における社会問題の階層構造と, そこに紐づけられた日本語 Wikipedia 記事を用いて訓練コーパスを作成する. 渡辺ら [2] が Wikipedia のカテゴリー階層を用いて構築した社会問題オントロジーは, 社会問題以外の Wikipedia 記事がノイズとして混じる問題があった. しかし, Wikidata の “subclass of” (wdt:P279) プロパティであれば, そのようなノイズが混じる危険性ははるかに小さい. よって, Wikidata の社会問題エンティティ (wd:Q1920219) のサブクラスを辿って得られたエンティティを社会問題タグとして用いる.

3 社会問題自動タグ付け手法

3.1 社会問題自動タグ付け手法

マルチラベル分類のための 4 つのアルゴリズム (BERT, Random Forest, Naive Bayes, K-Nearest Neighbor) を比較し, Web 記事を自動的

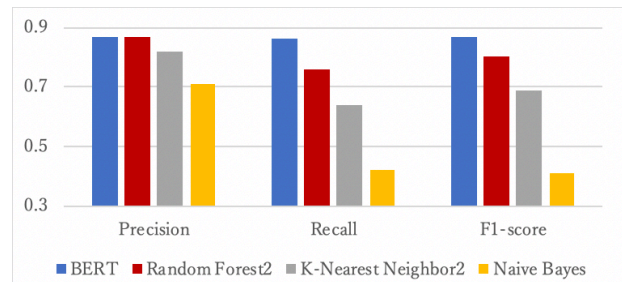


図 1. 各アルゴリズムの分類精度の比較

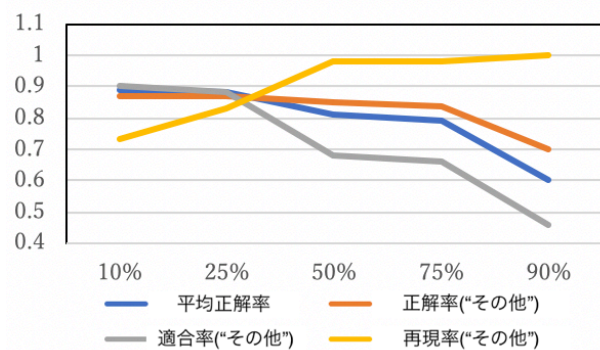


図 2. 負事例の割合を変化させたときの精度の推移

にタグ付けして分類する. モデルの精度を比較するために, 作成した社会問題コーパスの第二階層である “犯罪” クラスを使用する. 図 1 に各分類モデルの分類精度の比較を示す. BERT が最も良く F 値, 再現率, 適合率全てで 0.8 を上回った. また, マクネマー検定 [3] を用いて BERT モデルと他のモデルとの間に統計的な有意差があるかどうかを検証した. 検証の結果, いずれの場合も有意水準 0.01 で有意差が認められた.

3.2 社会問題自動タグ付け手法

社会問題記事を高精度に分類するためには, 社会問題について記述していない負事例を学習データに含める必要がある. そのために, 日本語版 Wikipedia から無作為に記事を抽出し, 社会問題を記述していない記事を “その他” としてラベリングすることで負事例を作成した. 図 2 は, 学習データに含まれる負事例の割合を変化させた場合の, 第二階層 “犯罪” に対する BERT 自動タ

Prototyping a Method for Social Issue Tagging to Web
Articled using BERT for Supporting Public Collaboration

[†] Tokutaka Hasegawa, Shun Shiramatsu

[‡] Nagoya Institute of Technology

グ付けモデルの予測精度を示している。予測モデルには、precisionとrecallの間でトレードオフの関係があり、負事例の割合が低い場合はprecisionが高く、反対に“その他”の割合が高い場合recallが高くなる。グラフにより、“その他”の割合が25%~30%のときに最適な分類であることがわかる。以上により、BERTによる自動タグ付けモデルでは、第一階層と第二階層のすべてのサブクラスに対して、負事例である“その他”を25%で学習させた。

3.1 社会問題自動タグ付け手法

本研究では、作成した社会問題コーパスに対して以下の3階層の分類タスクを設定した。これらの分類には汎用言語表現モデルBERTを用いた。また、事前学習モデルと文のトークン化に使用したSentencePieceも、日本語版Wikipediaのデータを用いて学習した。

1. 第ゼロ階層：与えられた文が社会問題についてのものか否かを判断する階層。これは、第一階層や第二階層に含まれない社会問題に関する記事を抽出するために必要である。
2. 第一階層：Wikidataにおける社会問題クラスにおいて特定のサブクラス（「犯罪」「災害」「虐待」「差別」「戦争」「疑似科学」「嫌がらせ」「環境問題」「迫害」）を選択し、関連する日本語版Wikipediaの記事名を分類する階層。
3. 第二階層：社会問題コーパスにおける第一階層のクラスにおいて特定のサブクラスを選択し、関連する日本語版Wikipediaの記事名を分類する階層。（例：「犯罪」に対して「テロ」「詐欺」「殺人」など）

4. 実験結果

表1は、第2節で説明した社会問題コーパスにおけるサブクラスにおいて、例として第一階層について、分類モデルの学習に使用しなかったデータを用いて分類精度の検証を行った結果を示している。図3は、同じく第一階層について、人手で収集した社会問題について書かれたニュース記事を用いて、分類精度の検証を行った結果の混同行列を示している。表1より、社会問題コーパスからの結果は、データ数の少ない“迫害”クラスを除き、比較的高い分類精度が得られていることがわかる。しかし、実際のニュース記事を用いた検証では、ほとんどのクラスで精度が低下しており、混同行列を確認すると分類モデルが多くのクラスを“その他”と誤分類していることがわかる。また、“戦争”や“迫害”のクラスの多くが“犯罪”クラスに分類されてい

表1. コーパスの第一階層における分類精度

社会問題				
	Precision	Recall	F1-score	Support
犯罪	0.93	0.76	0.84	2571
災害	0.97	0.82	0.89	1773
虐待	0.83	0.73	0.77	248
差別	0.89	0.75	0.81	255
戦争	0.91	0.82	0.86	667
疑似科学	0.94	0.87	0.91	264
嫌がらせ	0.87	0.75	0.80	166
環境問題	0.92	0.82	0.86	283
迫害	0.79	0.45	0.58	42
その他	0.74	0.97	0.84	3090
Accuracy			0.85	9359

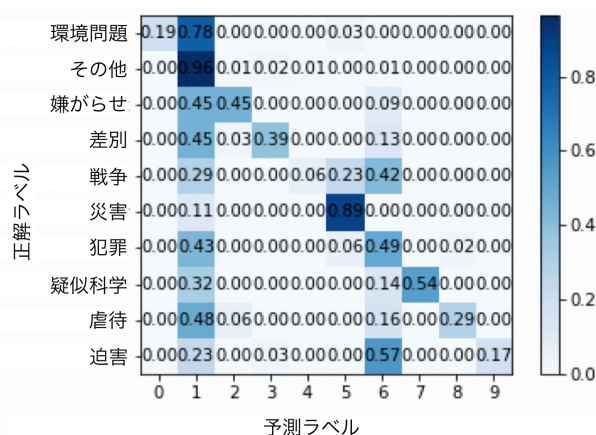


図3. Web記事における分類の混同行列

るが、これは完全に誤っているとは言えず、1つの文章に複数のタグを付与する必要があると考えられる。

5. おわりに

本研究では、社会問題に関するWeb記事の自動タグ付けシステムを試作した。実験の結果、社会問題コーパスにおけるタグ付け精度はF値0.85以上が得られ、実際のWeb記事ではF値0.60が得られた。今後は、予測精度向上のためクラウドソーシングを用いたデータ収集や探索的閲覧インターフェイスを開発したい。

謝辞 本研究の一部は、NEDO (JPNP20006), JST CREST (JPMJCR15E1, JPMJCR20D1), および科研費 (17K00461) の支援を受けた。

参考文献

[1] Devlin, Jacob, et al. "BERT: Pre-training of deep bidirectional transformers for language understanding." arXiv preprint arXiv:1810.04805 (2018).
 [2] Watanabe, M., et al. "Tag-based Approaches to Sharing Background Information regarding Social Problems towards Facilitating Public Collaboration." Proc. of eGose '17, 2017.
 [3] McNemar, Q. "Note on the sampling error of the difference between correlated proportions or percentages." Psychometrika 12.2 (1947): 153-157.