

マルチラベル付き日本語レビュー文章の分類

1

ソフトウェアシステム研究グループ

M1 楠本祐暉

発表の流れ

- 1. はじめに
- 2. 要素技術
- 3. データセット
- 4. 提案手法
- 5. 実験
- 6. まとめと今後の課題

発表の流れ

- ➡ 1. はじめに
- ➡ 2. 要素技術
- ➡ 3. データセット
- ➡ 4. 提案手法
- ➡ 5. 実験
- ➡ 6. まとめと今後の課題

はじめに

- ・ 近年、説明可能な人工知能が注目されている



自然言語処理の分野では…

- ・ 文章のセンチメントタグがなぜポジティブなのか、ネガティブなのかの説明が求められている

関連研究

- ・ アスペクトベースの感情分析

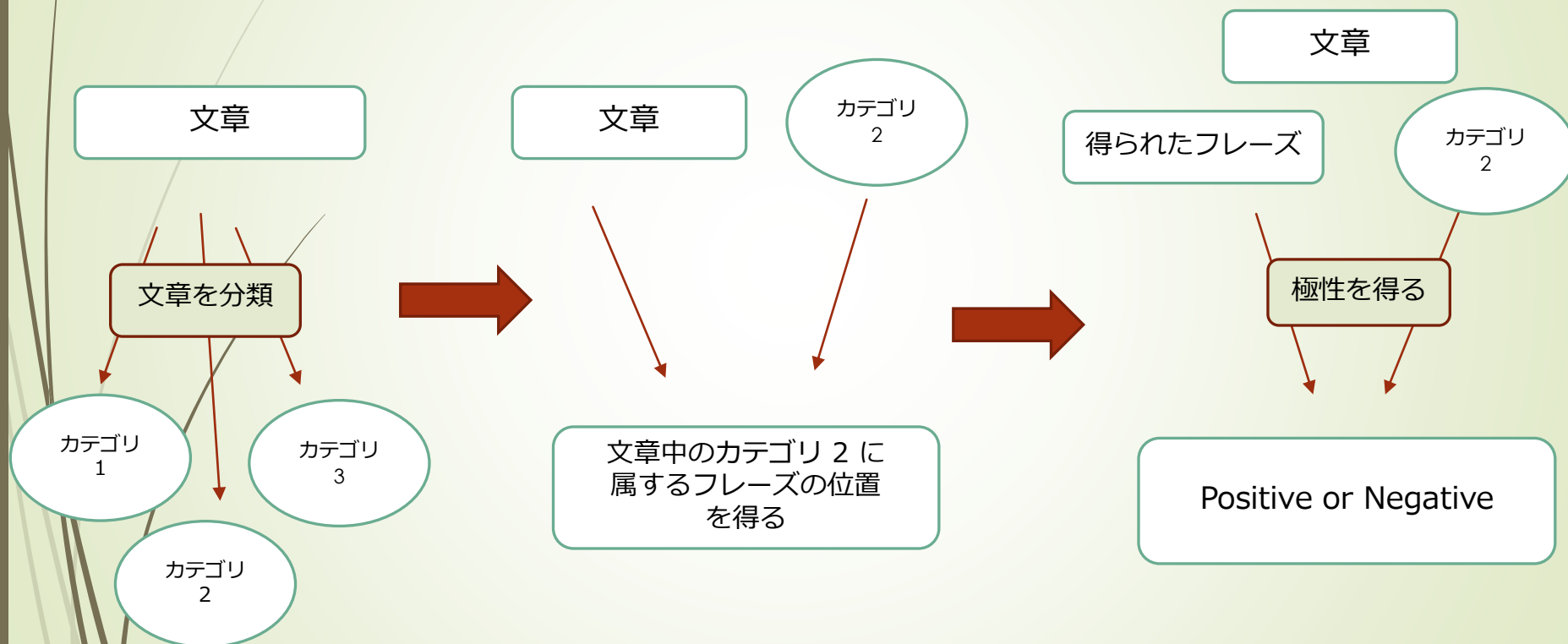
文章中に含まれるアスペクト情報を利用



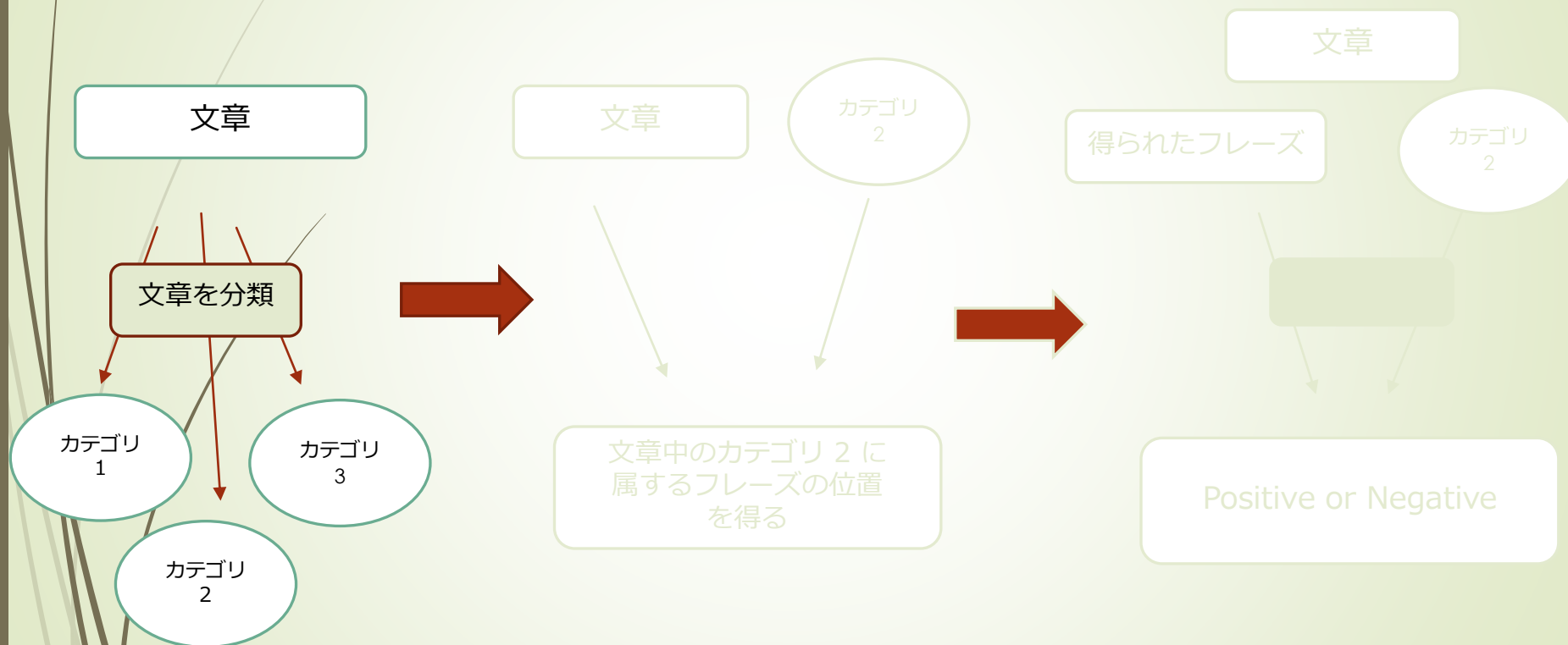
- ・ 文章が何を対象？
- ・ その対象の属性

どのようなことについて書かれた文章なのかを分析

アスペクトベースの感情分析とは



本研究では…



発表の流れ

- ➡ 1. はじめに
- ➡ 2. 要素技術
- ➡ 3. データセット
- ➡ 4. 提案手法
- ➡ 5. 実験
- ➡ 6. まとめと今後の課題

BERT (Bidirectional Encoder Representations from Transformers)

- 複数の双方向 Transformer に基づく汎用言語モデル
- 文章に依存した各単語、および文章の分散表現が得られる
- 本研究では、日本語 Wikipedia を用いた事前学習済み BERT モデルを使用
(東北大学 乾・鈴木研究室)

発表の流れ

- ➡ 1. はじめに
- ➡ 2. 要素技術
- ➡ 3. データセット
- ➡ 4. 提案手法
- ➡ 5. 実験
- ➡ 6. まとめと今後の課題

データセット

- ▶ 楽天トラベルレビューの
アспектセンチメントタグ付きコーパスを使用
- ▶ 楽天トラベルのレビュー約7万文について、
立地、部屋、食事等の7項目のアспектに対する
ポジティブまたはネガティブのタグが
付与されている
- ▶ 文章があるクラスに属していれば 1、属していな
ければ 0 というラベルが付与されている

楽天データセット

https://dsc.repo.nii.ac.jp/?action=pages_view_main&active_action=repository_view_main_item_detail&item_id=1752&item_no=1&page_id=13&block_id=21

作成者

<https://global.rakuten.com/corp/>

データの具体例

- 属しているクラス数を k とし、クラス数を N とすると、クラス情報は k of N ベクトルの形式になる

テキスト	朝食 positive	夕食 Negative	夕食 positive	夕食 negative	風呂 positive	風呂 negative	サービス positive	サービス negative	立地 positive	立地 negative	設備 positive	設備 negative	部屋 positive	部屋 negative
お部屋も広くて, お料理もとても美味しく, 部屋の露天風呂からは星がプラネタリウム のように広がっていて, とにかく最高でした	1	0	1	0	1	0	1	0	0	0	0	0	1	0
一部の方が指摘した通り, 廊下が タバコ臭いのが気になりました	0	0	0	0	0	0	0	1	0	0	0	1	0	0
立地と値段で決めました.	0	0	0	0	0	0	0	0	1	0	0	0	0	0

データの具体例

この 12 クラスを使用した

テキスト	朝食 positive	夕食 Negative	夕食 positive	夕食 negative	風呂 positive	風呂 negative	サービス positive	サービス negative	立地 positive	立地 negative	設備 positive	設備 negative	部屋 positive	部屋 negative
お部屋も広くて, お料理もとても美味しく, 部屋の露天風呂からは星がプラネタリウムのように広がっていて, とにかく最高でした	1	0	1	0	1	0	1	0	0	0	0	0	1	0
一部の方が指摘した通り, 廊下がタバコ臭いのが気になりました	0	0	0	0	0	0	0	1	0	0	0	1	0	0
立地と値段で決めました.	0	0	0	0	0	0	0	0	1	0	0	0	0	0

発表の流れ

- ➡ 1. はじめに
- ➡ 2. 要素技術
- ➡ 3. データセット
- ➡ 4. 提案手法
- ➡ 5. 実験
- ➡ 6. まとめと今後の課題

提案手法

文章を多クラスに分類するモデルの提案

1. 分散表現を獲得
2. 分類器による推定



発表の流れ

- ➡ 1. はじめに
- ➡ 2. 要素技術
- ➡ 3. データセット
- ➡ 4. 提案手法
- ➡ 5. 実験
- ➡ 6. まとめと今後の課題

実験

- ▶ 立っているラベルをどれ程予測できるのかを調べる



- ▶ 全ラベルのうち、少なくとも 1 つのラベルが立っているデータ群のみを用いて実験

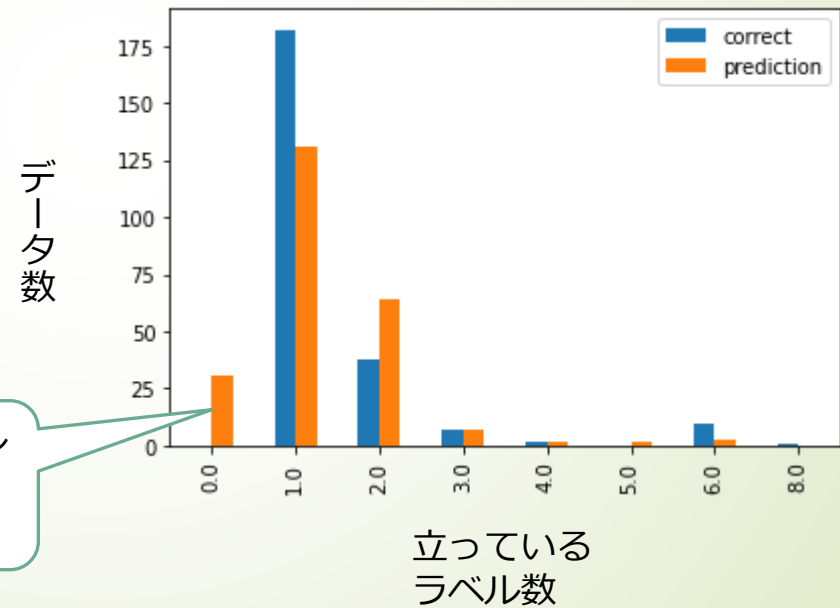
実験

➡ 予測方法

1. BERT の出力に対して閾値を設ける
2. 値が閾値を下回れば 0、上回れば 1 にする
3. 全ラベルが立たないデータに対しては、
最大値のみを 1 にする

実験

- ➡ 3 の処理を行わないと...



少なくとも1つのラベル
が立たないと予測結果
として不適切

識別機の評価指標

➡ 精度評価

正解率 (Accuracy), 適合率 (Precision),
再現率 (Recall), F1 値 (F1) を用いる

		真の結果	
		Positive	Negative
予測結果	Positive	TP	FP
	Negative	FN	TN

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$F = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

識別器の評価指標

- 全データ中の 2 割をテストデータ、8 割を訓練データとした
- 訓練データを 5 分割し、5 分割検証をすることで 5 個の識別機を作成した
- このうち、最も正解率の高いモデルを用いてテストデータでの評価をした

実験時のパラメータ

パラメータ	値
入力層の次元数	768
出力層の次元数	12
バッチサイズ	12
目的関数	BCE with Logistic Loss
エポック数	20

識別の精度

	夕食 positive	夕食 negative	風呂 positive	風呂 negative	サービス positive	サービス negative	立地 positive	立地 negative	設備 positive	設備 negative	部屋 positive	部屋 negative
正解率	0.858 3	0.958 3	0.912 5	0.958 3	0.816 7	0.833 3	0.875 0	0.983 3	0.812 5	0.879 2	0.908 3	0.920 8

識別の精度（ラベルが立っている部分のみ）

	夕食 positive	夕食 negative	風呂 positive	風呂 negative	サービス positive	サービス negative	立地 positive	立地 negative	設備 positive	設備 negative	部屋 positive	部屋 negative
適合率	0.8182	0.25	0.5676	0.1818	0.6486	0.4324	0.7037	0	0.2895	0.4231	0.6829	0.3333
再現率	0.5806	0.3333	0.8076	0.6667	0.7273	0.4571	0.4634	0	0.3793	0.44	0.7568	0.1875
F 1 値	0.6792	0.2857	0.6667	0.2857	0.6857	0.4444	0.5588	Nan	0.3284	0.4313	0.7179	0.2400

識別結果

データ	データ数
全テストデータ	240
全ラベルが正解	85
一部ラベルが正解 (立っているクラスのみ)	59
全ラベルが不正解 (立っているクラスのみ)	96

発表の流れ

- ➡ 1. はじめに
- ➡ 2. 要素技術
- ➡ 3. データセット
- ➡ 4. 提案手法
- ➡ 5. 実験
- ➡ 6. まとめと今後の課題

まとめ

- 本研究ではアノテーションされたアスペクトベースのデータを用いて評判分析を行った
- すべてのクラスの正解率の平均は 0.8930 であった
クラス 1 のみの F1 値の平均値は 0.4437 であった
- ほとんどのデータにおいてラベルが立っていないクラスが多いため、予測が 0 に寄ってしまう問題があった

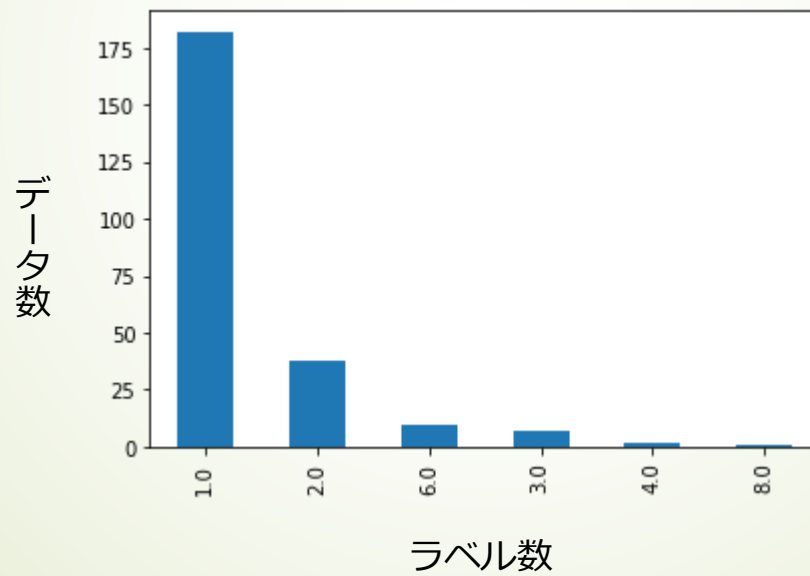
今後の課題

- 全体的な精度の向上とその手法の模索に取り組む
- 今回用いたデータセット以外での実験を試みる
- アスペクトベースの感情分析のサーベイ
- 一部正解データや全不正解データについて、
関連研究であるアスペクトベースの感情分析の手法
を用いて分析する

- 楽天グループ株式会社が国立情報学研究所を通じて研究者に提供しているデータセットです

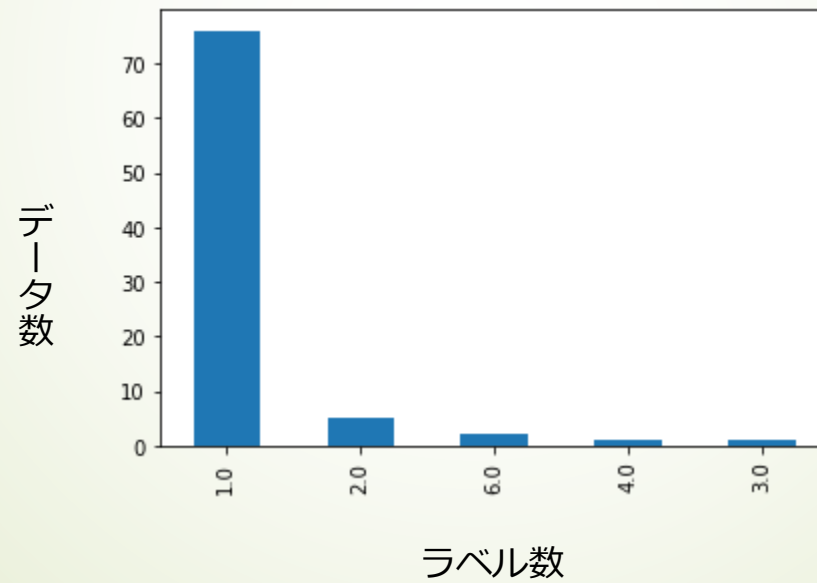
識別結果

- 全ラベルが正解したデータのラベル数の分布



識別結果

- 全ラベルが正解したデータのラベル数の分布



はじめに

- ・近年、説明可能な人工知能が注目されている



このような研究に取り組むために...

- 日本語レビュー文章を用いた多値分類に取り組む