

書き手の性格情報を用いた 感情強度推定

鈴木 陽也[†] 秋山 和輝[†] 梶原 智之[†] 二宮 崇[†]
武村 紀子[‡] 中島 悠太[‡] 長原 一[‡]

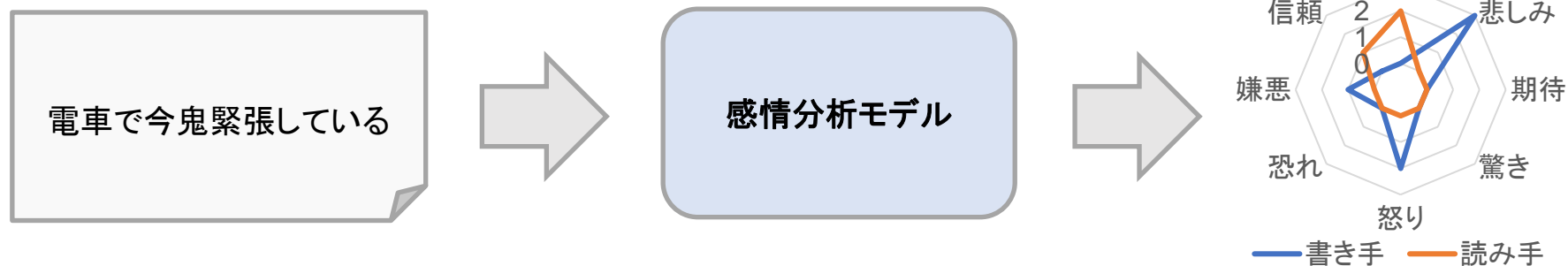
[†]愛媛大学 [‡]大阪大学

JSAI2022

はじめに

感情分析：

テキストから人間の感情強度を予測するタスク



感情分析の応用：

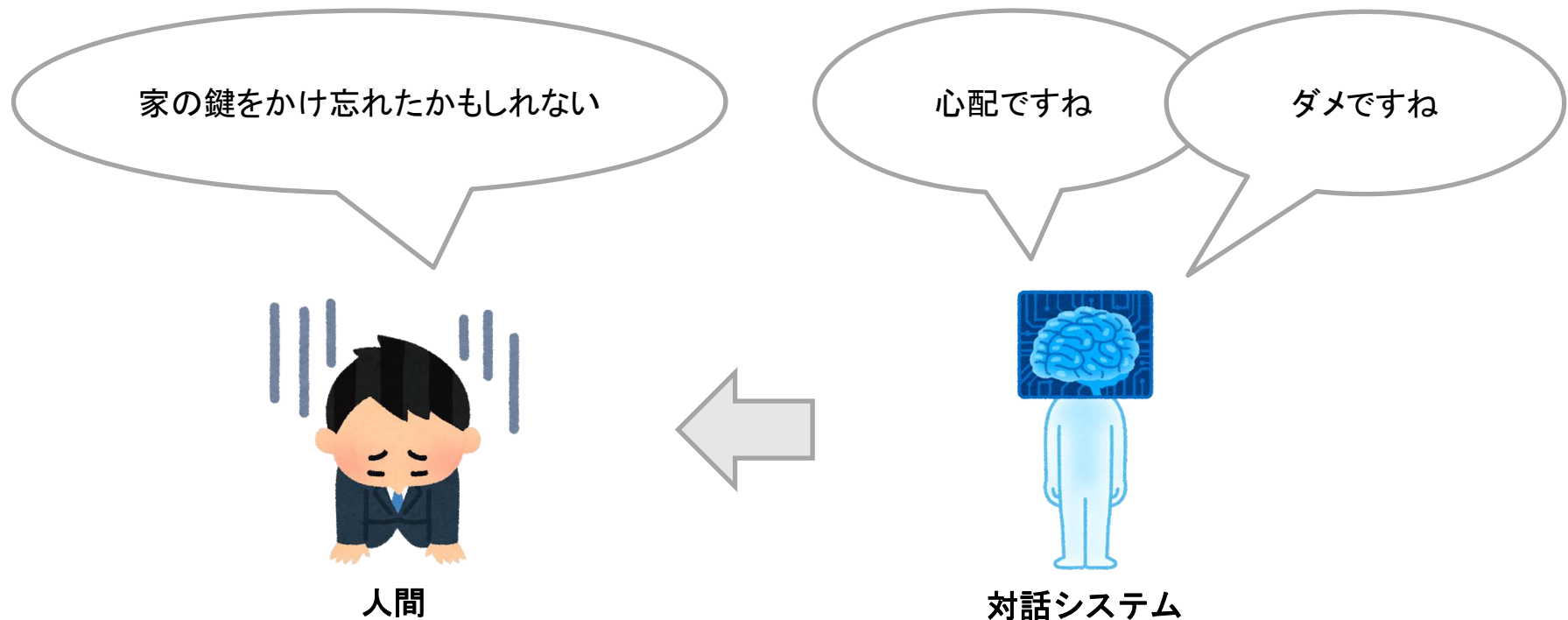
- 対話システム [Tokuhsa 08]
- ソーシャルメディアマイニング [Stiegliz and Dang-Xuan 13]

[Tokuhsa+ 08] **Emotion Classification Using Massive Examples Extracted from the Web**, in Proc of COLING, 2008

[Stiegliz and Dang-Xuan 13] **Emotions and Information Diffusion in Social Media — Sentiment of Microblogs and Sharing Behavior**, Journal of Management Information Systems, 2013

例：対話システムにおける感情分析の活用

- 積極的に話を聞いてくれるパートナーとしての役割
- 共感的な応答を得るために感情分析を利用



関連研究 (1/5)

感情極性の分類が主流

- ポジティブ、ニュートラル、ネガティブ
- Stanford Sentiment Treebank^[1]やSuzuki^[2]らのコーパスで利用



最近は、より詳細な感情強度の推定も利用

- Ekmanの基本6感情^[3]
- Plutchikの基本8感情^[4]

[1] Socher et al. **Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank**, In Proc of EMNLP, 2013

[2] Suzuki. **Filtering Method for Twitter Streaming Data Using Human-in-the-Loop Machine Learning**, Journal of Information Processing, 2019

[3] Ekman. **An Argument for Basic Emotions**, Cognition and Emotion, Cognition and Emotion, 1992

[4] Plutchik. **A General Psychoevolutionary Theory of Emotion**, Theories of Emotion, Theories of Emotion, 1980

関連研究 (2/5)

Ekmanの基本6感情^[3]

- 喜び、悲しみ、驚き、怒り、恐れ、嫌悪
- SemEval-2007の感情分析コンペ^[5]などで利用

Plutchikの基本8感情^[4]

- 喜び、悲しみ、期待、驚き、怒り、恐れ、嫌悪、信頼
- WRIME^[6]などで利用



本研究では、日本語コーパスの**WRIME**^[6]を用いる

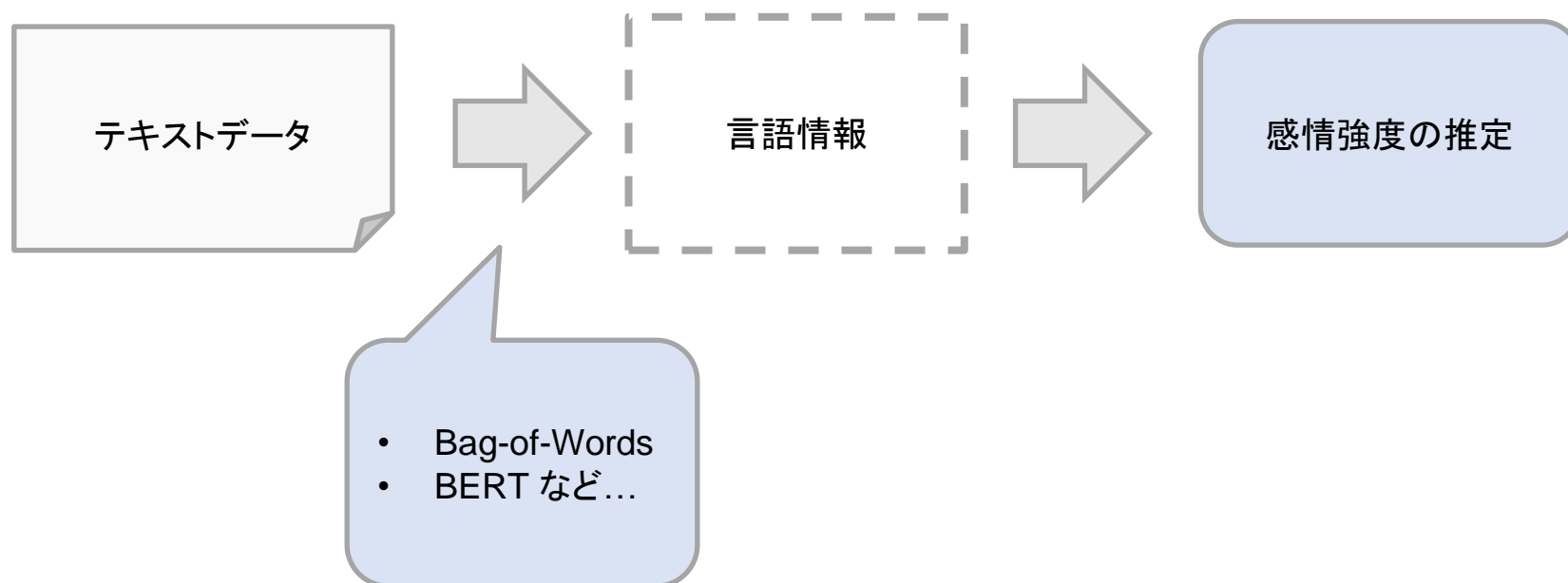
[5] Strapparava and Mihalcea. **SemEval - 2007 Task 14: Affective Text**, In Proc of SemEval, 2013

[6] Kajiwara et al.. **WRIME: A New Dataset for Emotional Intensity Estimation Subjective and Objective Annotations**, In Proc of NAACL, 2021

関連研究 (3/5)

一般的な感情分析モデル：

- テキストデータから言語情報を抽出
- 感情強度（喜びや悲しみなど...）を推定



関連研究 (4/5)

Bag-of-Words :

- ある単語が出現する頻度を言語情報とする
- 文脈を考慮できていない

①この店のアイスはおいしいです。

②店員の態度が悪いのが残念です。

③この店のアイスはおいしいです。ただ、店員の態度が悪いのが残念です。

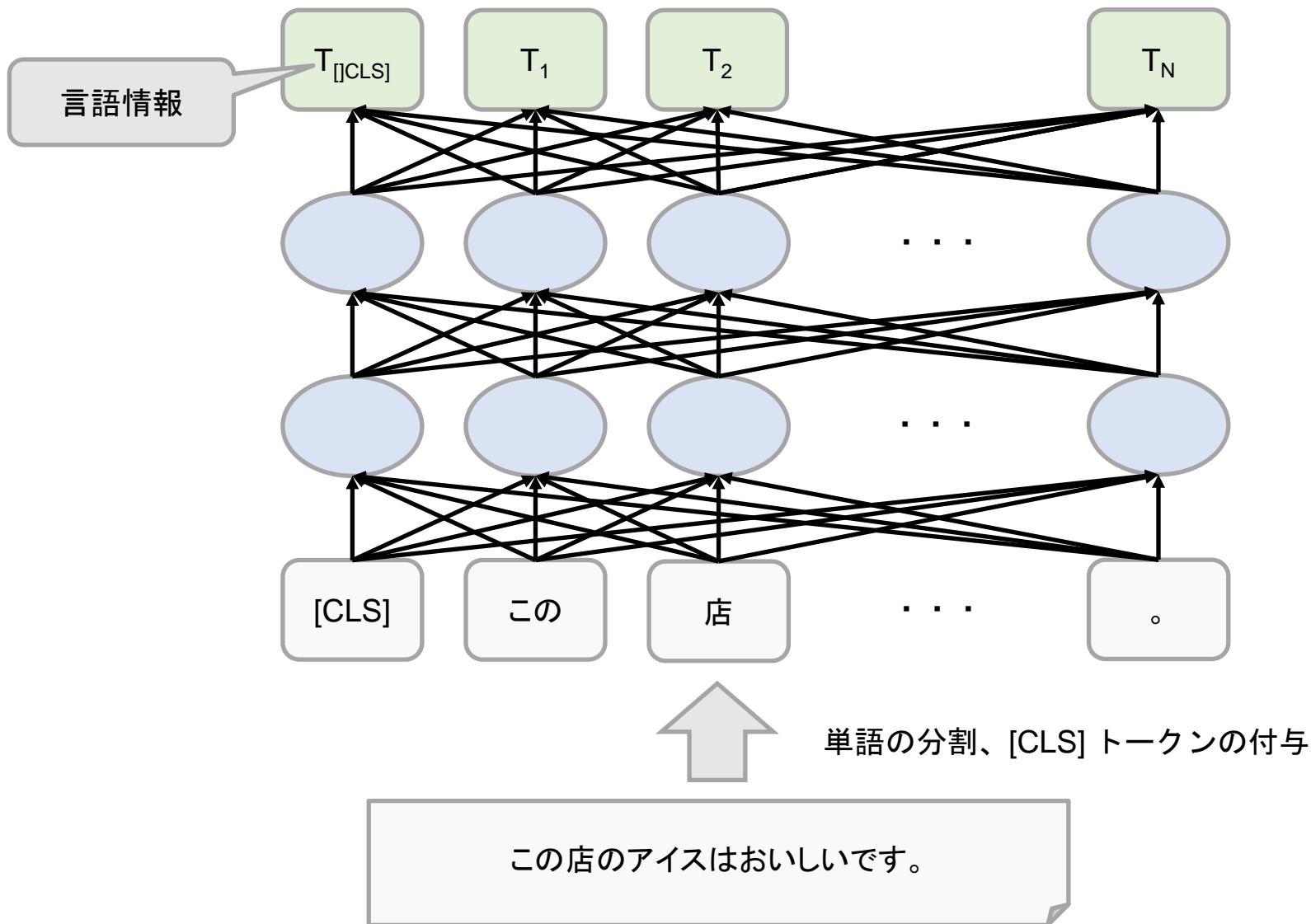


単語分割、単語の出現頻度をカウント

	この	店	店員	アイス	態度	おいしい	悪い	残念	です
①	1	1	0	1	0	1	0	0	1
②	0	0	1	0	1	0	1	1	1
③	1	1	1	1	1	1	1	1	2

関連研究 (5/5)

BERT : 文脈を考慮した言語情報を生成するモデル

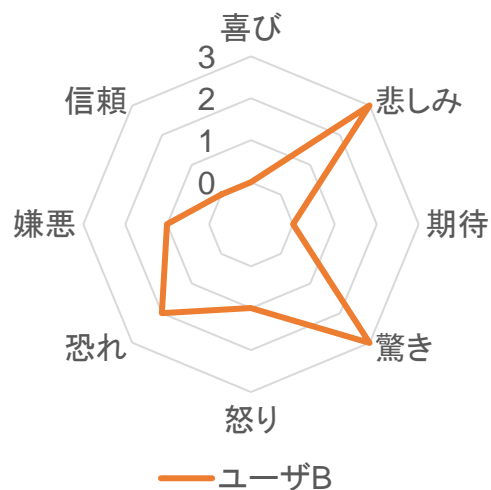
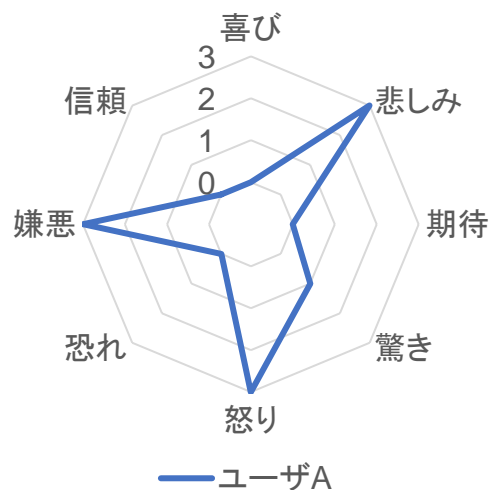


研究課題

同じテキストを書いたとしても、
背景にある各ユーザの感情はそれぞれ異なりうる

⇒ユーザ固有の情報を考慮した感情分析が重要

車のタイヤがパンクしてた。。いたずらの可能性が高いんだって。。



性格診断（1/3）

Big Five^[9] :

- 人間の性格が5因子の強度によって決まる
- 外向性、情緒不安定性、開放性、誠実性、調和性

Big Five尺度の標準化^[10] :

- 以下の二つの要素を対応付ける

60項目の性格形容語

- 陽気な
- 話し好きな
- 悩みがち
- 独創的な
- いい加減な など



Big Five 5因子

- 外向性
- 情緒不安定性
- 開放性
- 誠実性
- 調和性

- e.g. 外向性 = 陽気な度合い + 話し好きな度合い + . . .

[9] Goldberg. **The Development of Markers for the Big-Five Factor Structure**, Psychological Assessment, 1992

[10] Saito et al. **Standardization of Big Five Scales Using the Adjective Check List.**, Kyushu University Psychological Research, 2001

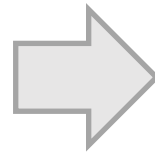
性格診断 (2/3)

ユーザごとに60項目のアンケートからBig Five 5因子を評価

アンケートの質問(60項目)	1~7
話し好き	2
悩みがち	6
いい加減な	4
温和な	4
無口な	6
不安になりやすい	6

⋮

1: 非常にあてはまらない
2: あまりあてはまらない
3: どちらかといえばあてはまらない
4: どちらともいえない
5: どちらかといえばあてはまる
6: わりとあてはまる
7: 非常にあてはまる



診断結果(5項目)	
外向性	60
情緒不安定性	38
開放性	58
誠実性	57
調和性	64

60項目の足し引きにより自動評価

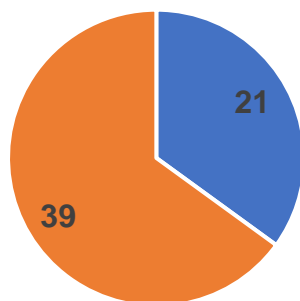
外向性 = 陽気な + 話し好きな + ... - 無口な + 48

性格診断 (3/3)

本研究で用いる感情分析データセットの対象者60人をクラウドソーシングで雇用

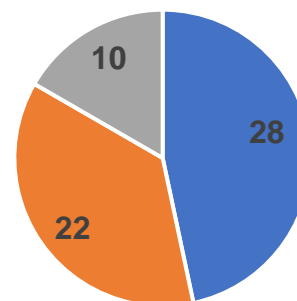
対象者：テキストの書き手かつ感情強度のアノテータ

男女別対象者の内訳



■ 男性 ■ 女性

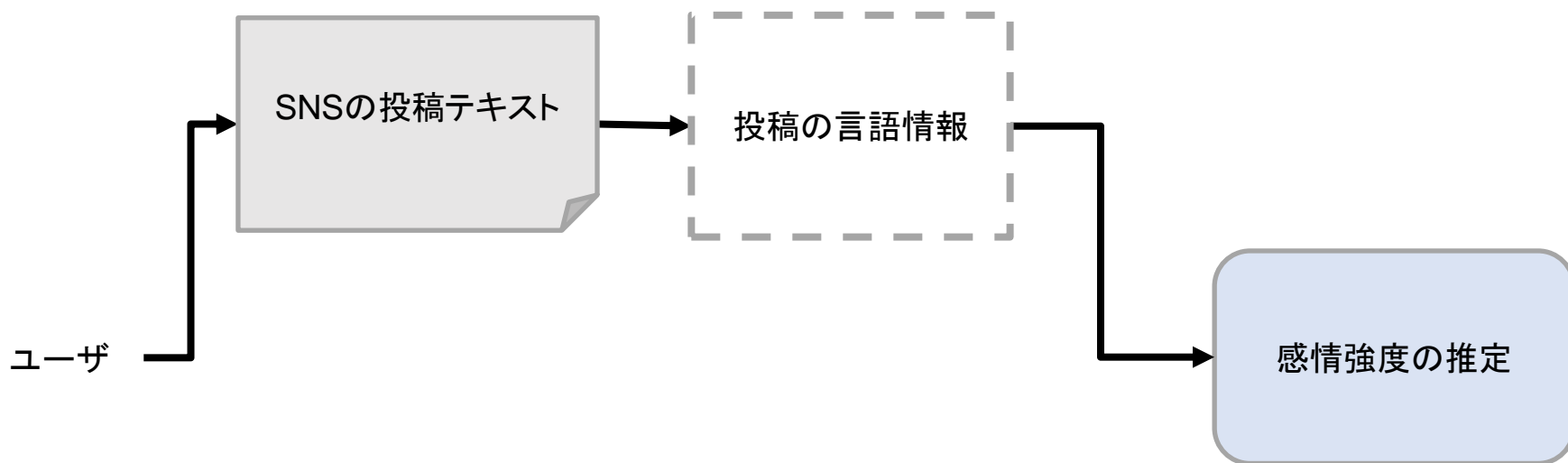
年代別対象者の内訳



■ 20代 ■ 30代 ■ 40代～

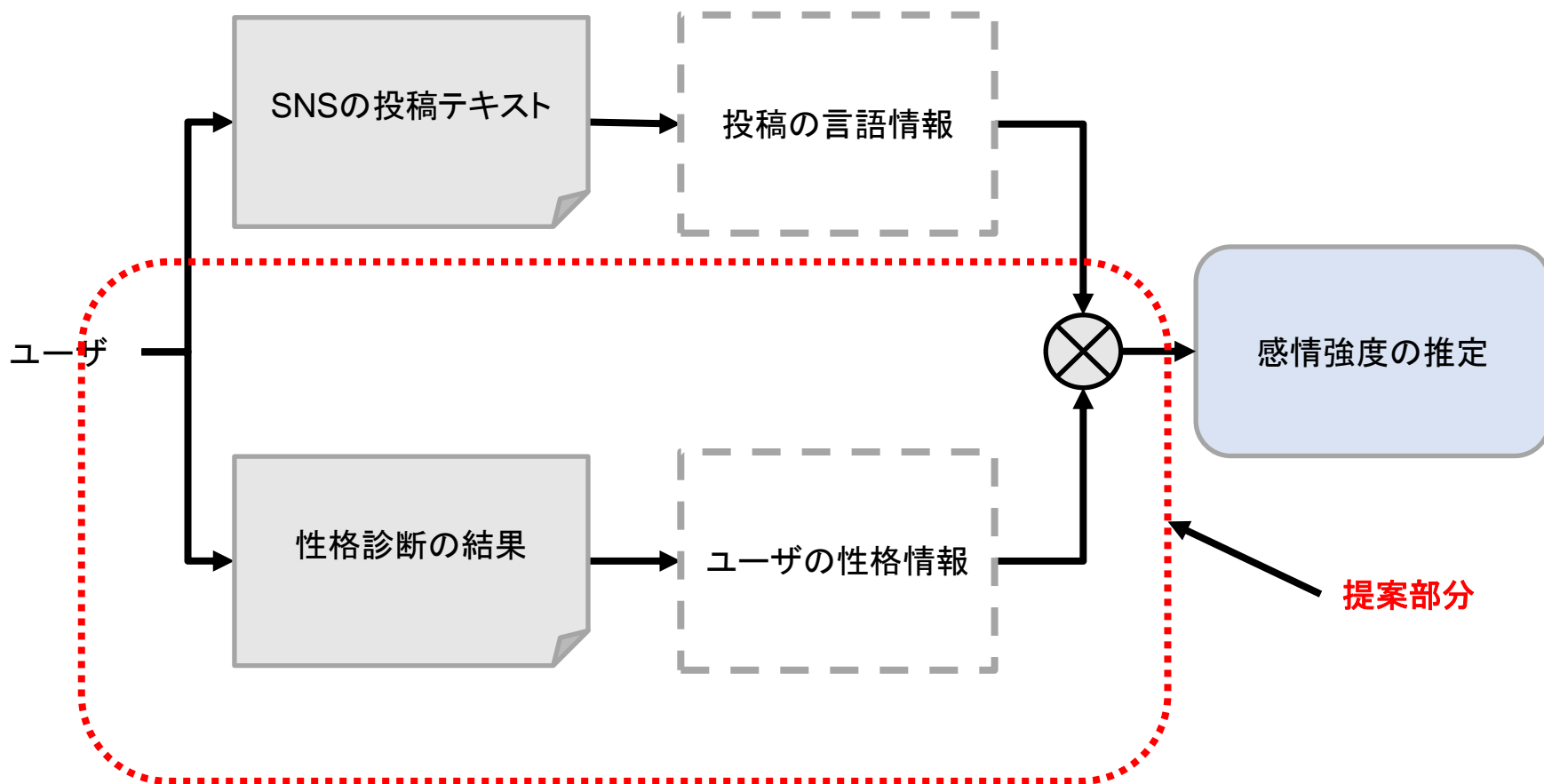
提案手法の概要

従来の感情分析モデルにユーザの性格情報を追加、
書き手に特化したモデルを提案



提案手法の概要

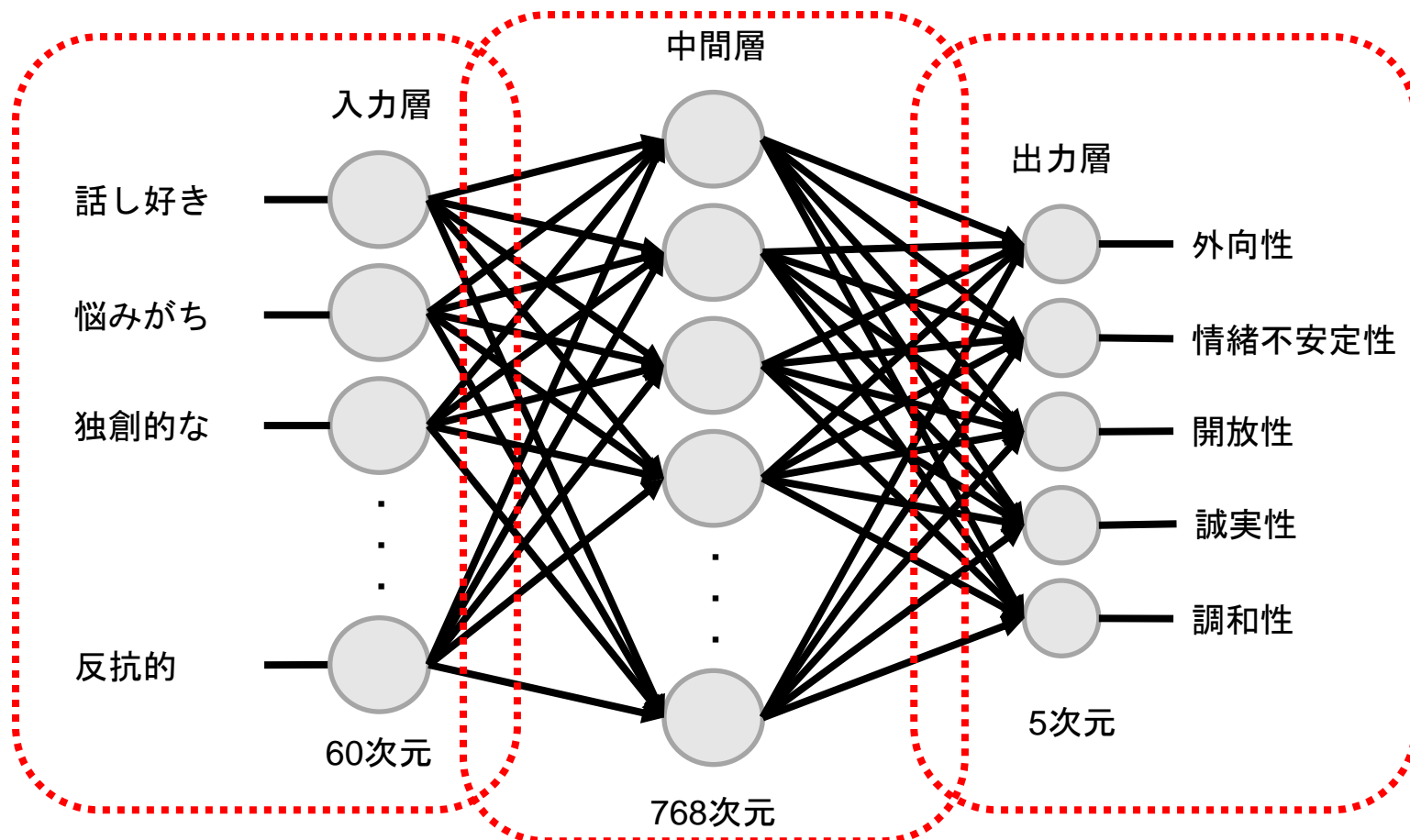
従来の感情分析モデルにユーザの性格情報を追加、
書き手に特化したモデルを提案



ユーザの性格情報

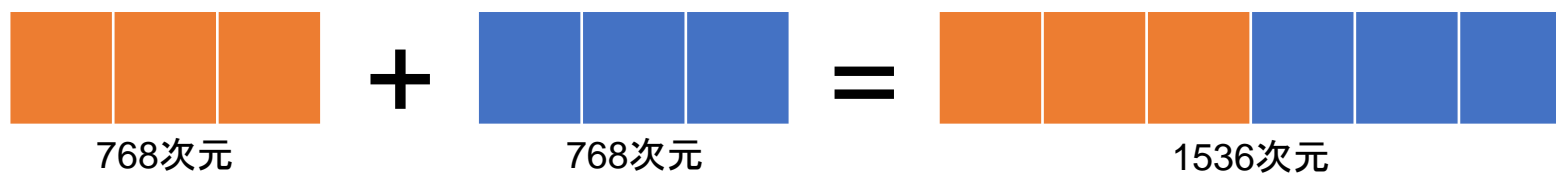
アンケート項目（60項目）から診断結果（5項目）
を予測する回帰タスクを学習

⇒ 中間層から特徴ベクトルを抽出



言語情報 × 性格情報

Concat法：投稿の言語情報とユーザの性格情報を結合



Product法：投稿の言語情報とユーザの性格情報を要素積



感情強度の推定実験

データセット：

- WRIME [6]
- 性格診断

モデル：

- BoW
- BERT (Wikipedia)
- BERT (SNS)
- 提案手法
 - BERT Concat
 - BERT Product

評価指標

- Quadratic Weighted Kappa (QWK)

	訓練用	検証用	評価用
人数	60	10	10
件数	40,000	1,200	2,000

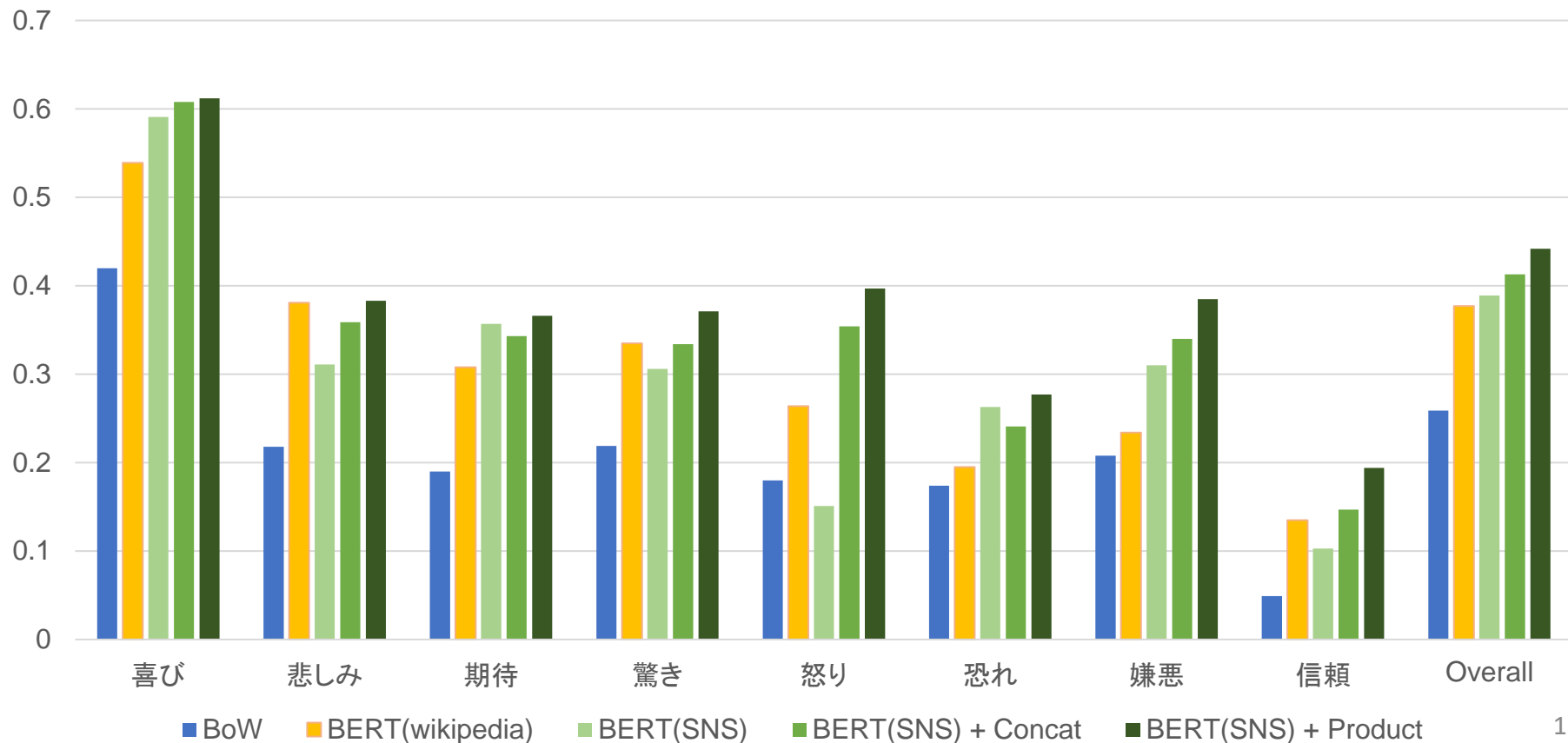
ベースラインモデルと提案手法の比較

総合評価Overallの結果より

提案手法 (Product法) が最高性能を達成

⇒ 性格情報の考慮が感情強度の推定に有効

[QWK]

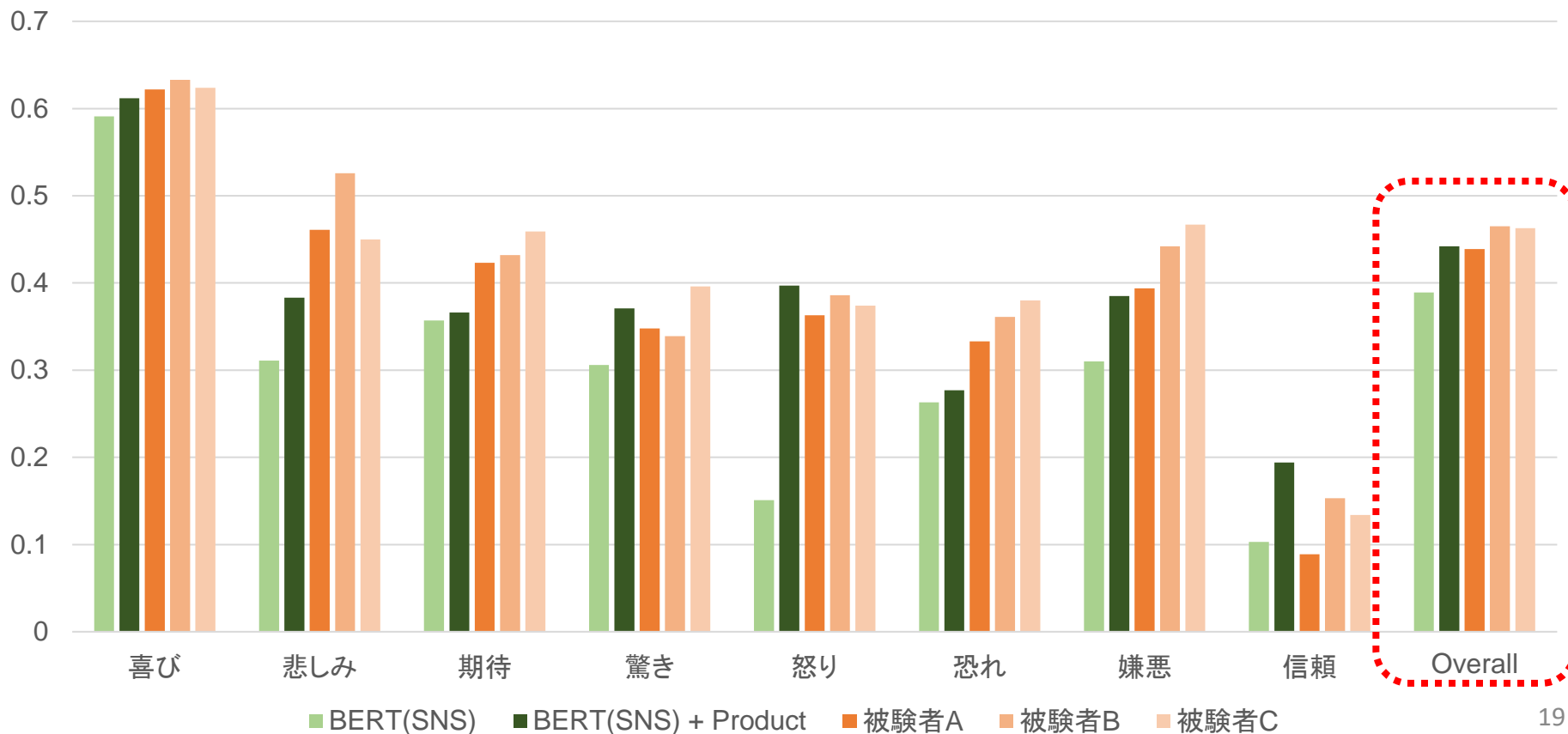


提案手法と人手評価の比較

総合評価Overallの結果より

提案手法は、被験者Aと同等の性能を達成

[QWK]



提案手法による性格考慮の例

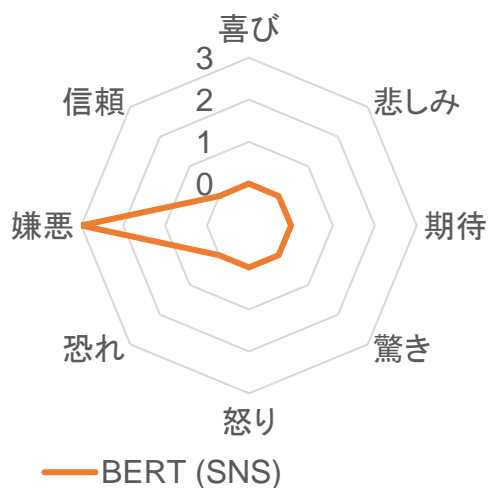
性格診断より、このユーザは怒りっぽい性格の持ち主であることがわかっている

大学の大声で馬鹿騒ぎするノリが嫌いすぎて学校行くの嫌になる

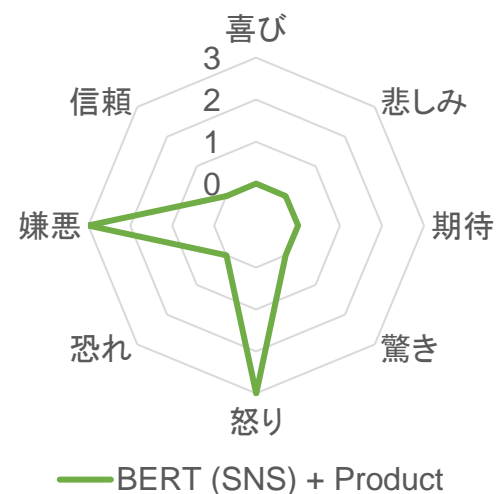
書き手の実際の感情



モデルが推定した書き手の感情



ベースラインモデル



提案モデル

書き手の性格情報を用いた感情強度推定

背景：大量のSNSデータにより感情分析の機会が増加

課題：同じテキストを書いたとしても、

背景にある各ユーザの感情はそれぞれ異なりうる

手法：ユーザの性格情報を考慮した感情分析モデルを提案

結果：提案手法がベースモデルの性能を上回り最高性能を達成

⇒書き手の感情推定における有効性が示された

