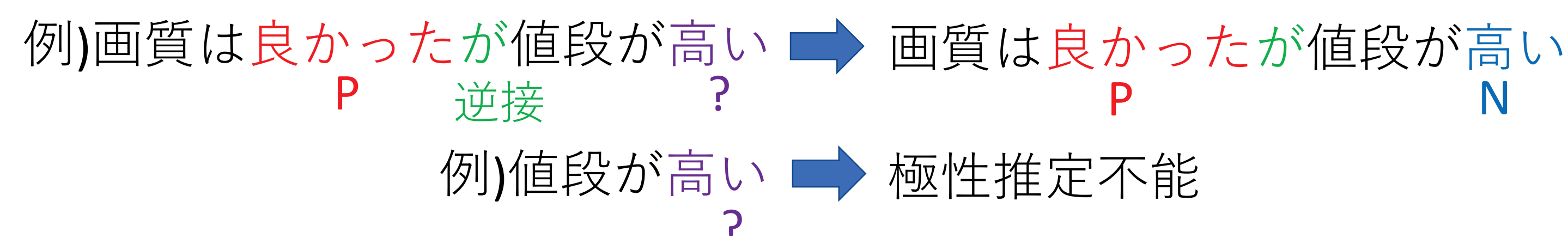


## 研究背景

・印象極性推定 P:ポジティブ N:ネガティブ ?:極性不明

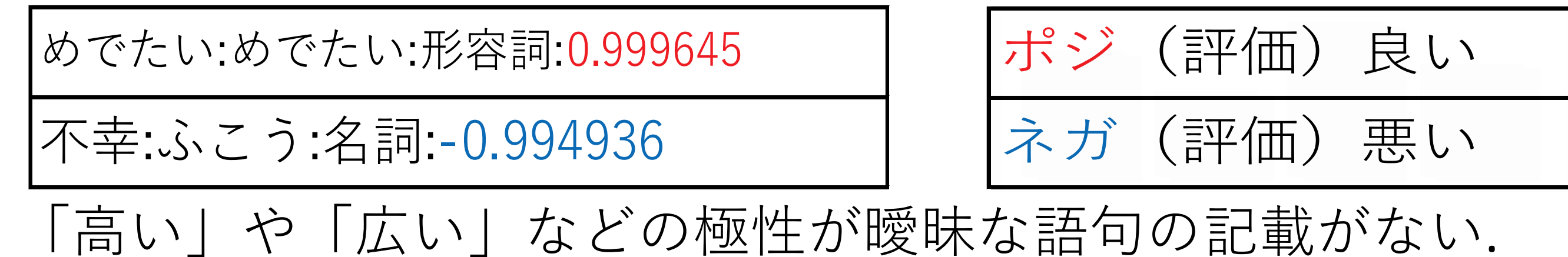
印象極性辞書による極性推定 文脈による極性推定[1]



・既存の印象極性辞書の問題点

単語感情極性対応表[2]

日本語評価極性辞書(用言編)[3]



## 研究目的

・印象極性の曖昧性

危険性が高い:N  
安全性が高い:P 高い:?

危険性が高い:N  
安全性が高い:P

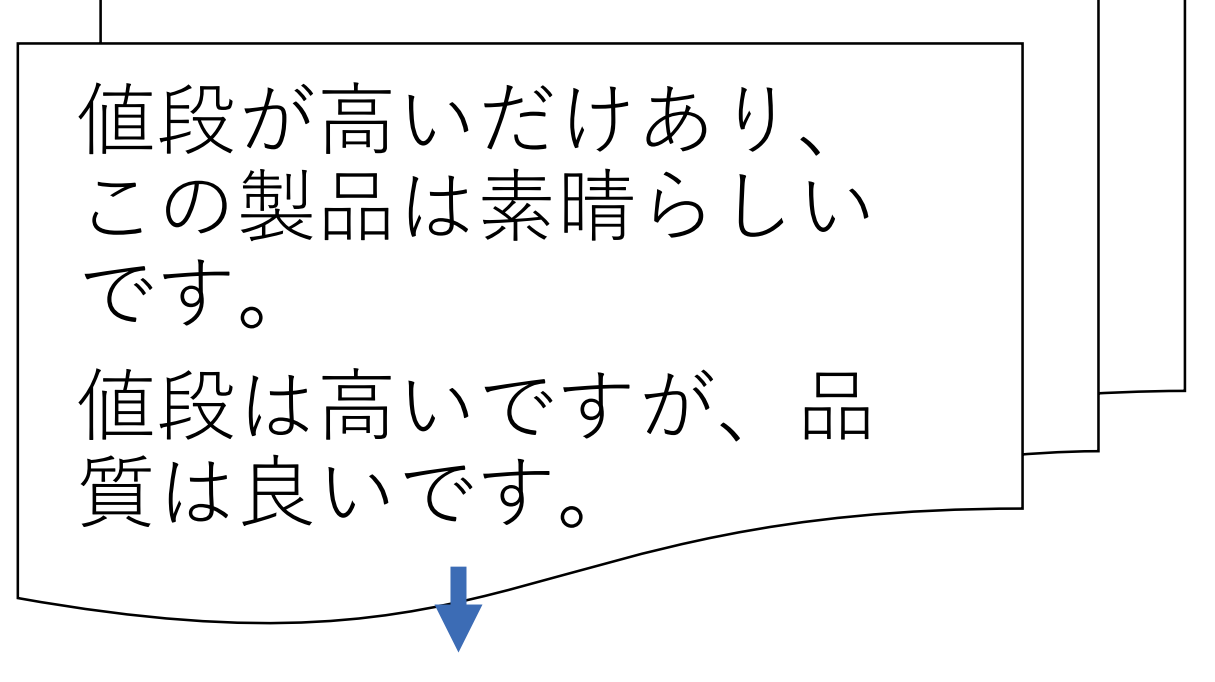
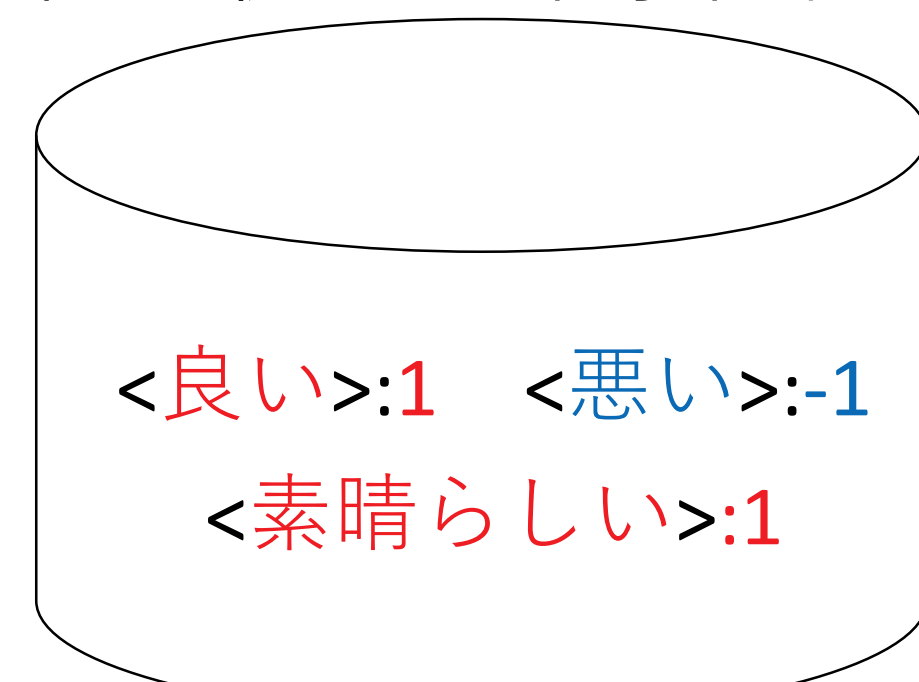
・文章の商品に対する評価情報[3]

評価情報を半自動抽出するため評価情報<対象,属性,評価表現>と定義。

<属性,評価表現>組の印象極性辞書を作成

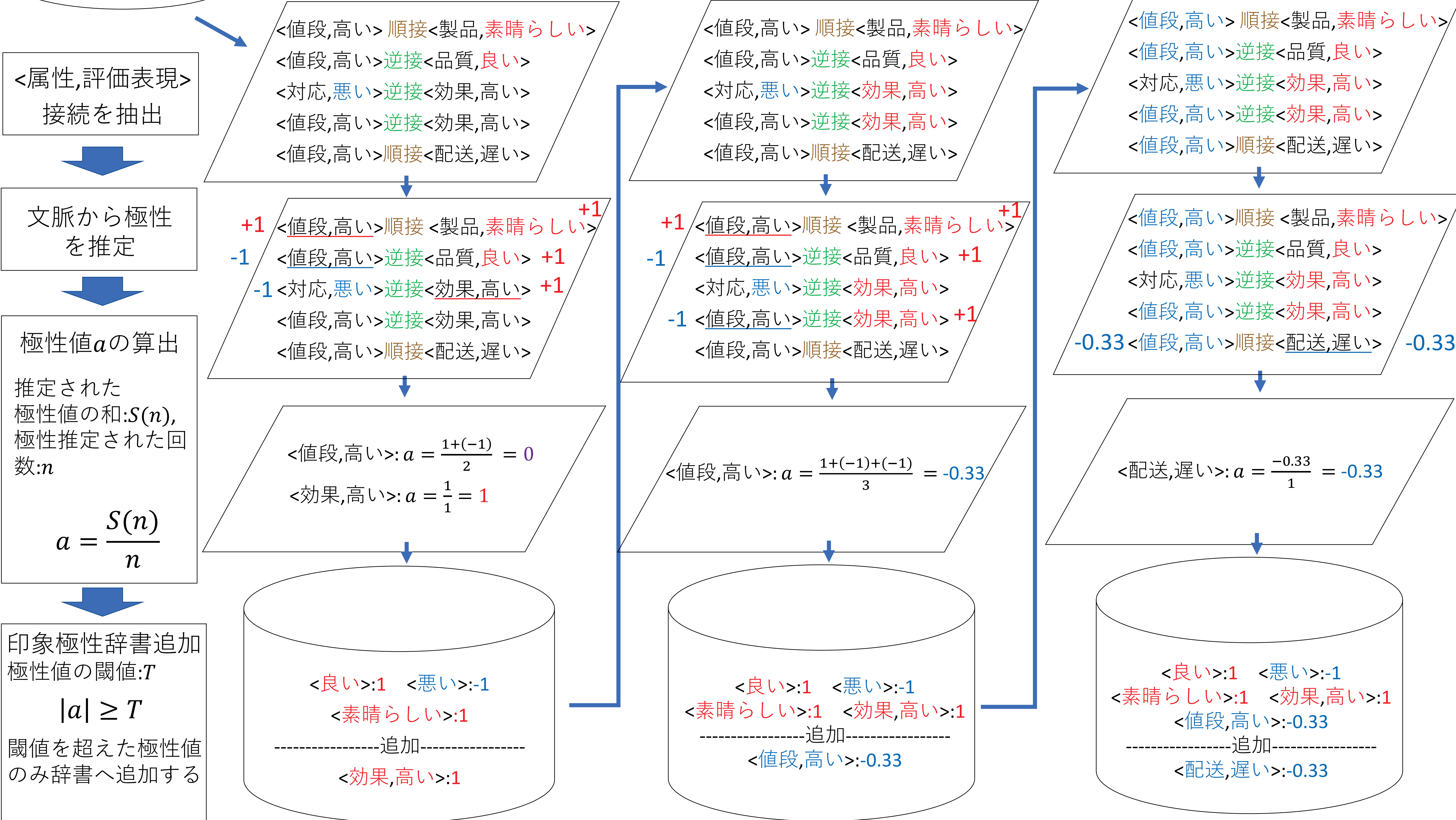
## 提案手法

初期状態の印象極性辞書



・使用するデータ

入力:楽天市場レビュー文章群(約7000万レビュー)[4]  
→前処理として文に分割  
初期状態の印象極性辞書:日本語評価極性辞書(用言編)  
→極性値をポジティブなら1,ネガティブなら-1として記載



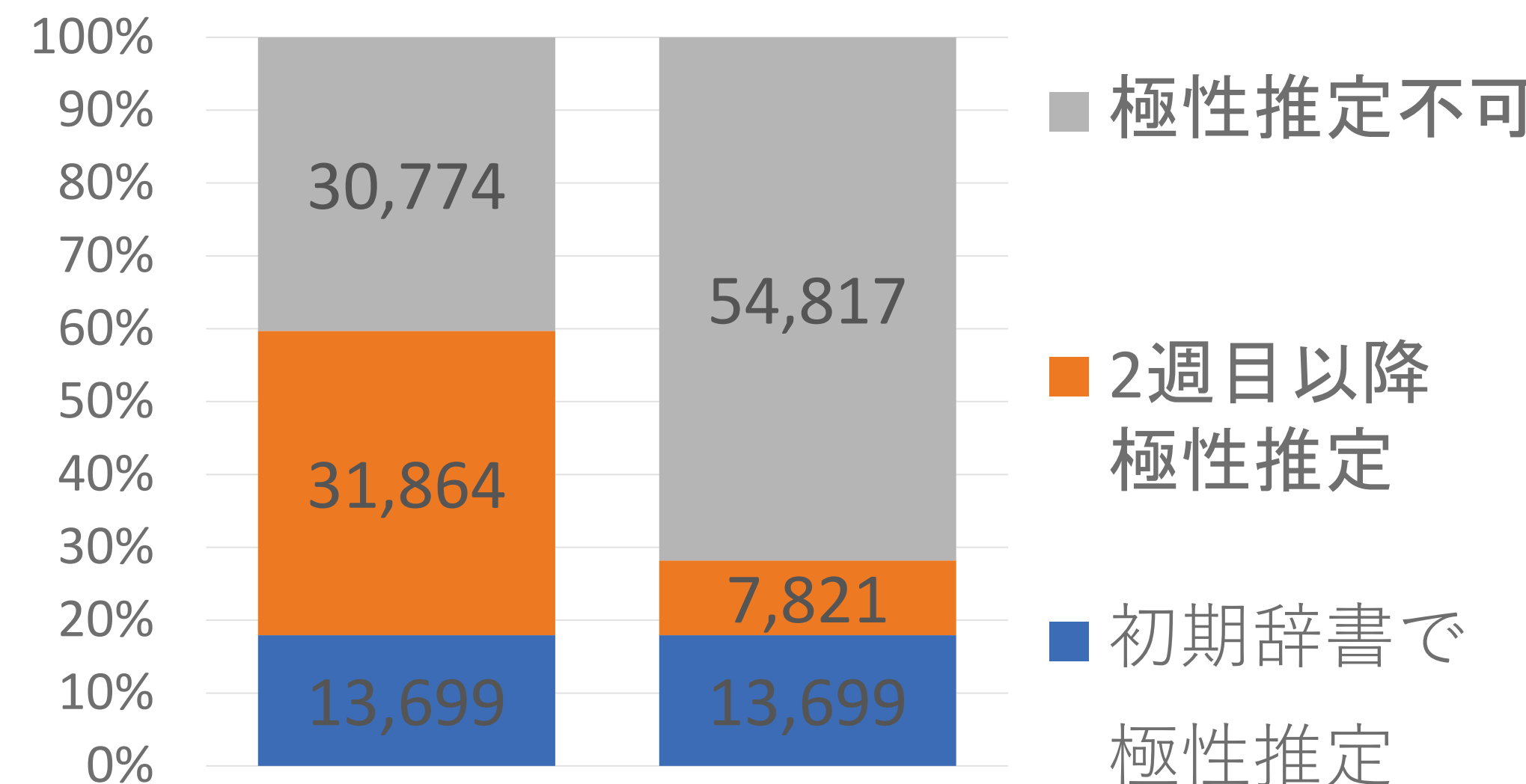
## 実験結果

・極性推定結果

レビュー文章:59,806件,レビュー文群:177,006文  
極性値閾値T:0.6

辞書に追加される極性の出現回数について閾値を定め比較した

・出現回数閾値なし・出現回数5回以上



・極性値と正解データの比較

出力結果	P	N	合計
正解データ			
P	37	0	37
N	8	1	9
合計	45	1	46

・極性推定ができないレビュー文の例  
ファスナーの開閉も容易で縫製も丁寧  
商品の納品が素早かった

・極性値算出に利用された評価情報の数と文章全体での出現数

評価情報	極性値算出に利用(回)	文章全体での出現数(回)
<発送,早い>	56	731
<配送,早い>	13	362
<梱包,丁寧>	12	267

<属性,評価表現>	出現回数	極性値
<発送,早い>	731	0.93
<配送,早い>	362	0.69
<梱包,丁寧>	267	1.00
<対応,早い>	235	0.99
<デザイン,可愛い>	184	0.77
<到着,早い>	176	1.00
<組み立て,簡単>	140	0.67
<違和感,ない>	138	0.75
<申し分,ない>	110	0.60
<色,可愛い>	103	0.60
<商品,ない>	92	0.79
<値段,手頃>	90	0.94
<取り付け,簡単>	86	1.00
<購入,よい>	75	0.65
<サイズ,ぴったり>	62	0.60
<対応,丁寧>	59	0.79
<値段,安い>	53	0.48
<商品,早い>	53	0.99
<発送,迅速>	50	1.00
<肌触り,よい>	47	0.93
<傷,ない>	46	0.60
<柄,可愛い>	45	1.00
<商品,綺麗>	44	0.98
<生地,柔らかい>	39	1.00
<注文,早い>	38	0.64
<変化,ない>	36	-0.60
<不満,ない>	36	0.86
<他,ない>	36	0.84
<梱包,綺麗>	34	0.85
<包装,丁寧>	30	0.92
<機能,ない>	28	0.87
<乾き,早い>	26	1.00
<種類,豊富>	25	0.95
<不備,ない>	23	0.81
<生地,よい>	23	0.68
<自体,ない>	18	0.79
<価格,ない>	17	0.68
<デザイン,ない>	17	0.76
<価格,安い>	16	0.48
<サイズ,大きい>	16	0.05
<発送,丁寧>	16	1.00
<配送,丁寧>	16	1.00
<サイズ,ない>	15	0.04
<何,問題>	14	0.55
<生地,薄い>	14	0.00
<品質,ない>	13	0.72
<商品,丁寧>	12	0.99
<チェーン,長い>	12	0.60
<見た目,可愛い>	12	0.47
<値段,高い>	12	-0.17

## 参考文献

[1] 漆原駿, 町田翔, 延澤志保, "印象極性による商品レビューの視覚化," 情報処理学会第 80 回全国大会, vol.2, no.7-01, pp.311-312, 2018.  
 [2] 高村大也, 乾孝司, 奥村学, "スピンモデルによる単語の感情極性抽出," 情報処理学会論文誌, Vol.47 No.02 pp. 627-637, 2006.  
 [3] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一. "意見抽出のための評価表現の収集" 自然言語処理Vol.12, No.3, pp.203-222, 2005.  
 [4] 楽天グループ株式会社 (2020): 楽天市場データ. 国立情報学研究所情報学研究データリポジトリ. (データセット). <https://doi.org/10.32130/idr.2.1>

## 謝辞

本研究では、国立情報学研究所のIDRデータセット提供サービスにより楽天グループ株式会社から提供を受けた「楽天データセット」([https://rit.rakuten.com/data\\_release/](https://rit.rakuten.com/data_release/))を利用しました。感謝申し上げます。