

疑似ラベルデータ作成に基づく楽天データセットの食品レビューからの観点抽出

S01

芦澤亜里紗, 瀧田慎, 大島裕明 (兵庫県立大学)

背景



- ・実物見えない
 - ・気軽に質問できない
- レビューを参考

目的



本研究

めんの**弾力**がもう少しあれば完璧

麺には**噛み応え**があり美味しいです

Rakuten

楽天データセット

Rakuten

楽天レビューデータ : 2015年から2019年
「食品」、「スイーツ・お菓子」 : 6,477,422件



観点



検索クエリ : こし ↔ **観点:硬軟**
レビュー : 弾力があります。

観点辞書作成



本「ことばは味を超える」 + 包装・梱包などの味表現以外の言葉
の味表現の分類項目

データ

a.硬軟

a. お肉がとても**柔らかい**。
b. 口に入れた瞬間とろける。



商品No : 42

①観点辞書データ

観点言葉
柔らかい
固い

観点
硬軟

②疑似ラベルデータ

同じ評価
同じ商品 のレビュー = aとbは **同じ観点**

学習

モデル

1. **Wikipedia**の言葉のみ知っているモデル



観点辞書データを用いた学習

a + 観点



モデル

観点を付けられるモデル

2. **Wikipedia + レビュー**の言葉を知っているモデル



疑似ラベルデータを用いた学習

a

b

モデル

文の類似がわかるモデル

評価

観点抽出器への入出力に対する**アンケート**

入力 : 歯が割れそう。
出力 : 硬軟

出力が適切か
○ or ×

再現率を
求める