

楽天市場の商品レビューを利用した評価表現辞書の学習と商品ジャンルに対応したレーダーチャート自動生成

電気通信大学大学院 情報理工学系研究科 金兵裕太 沼尾雅之

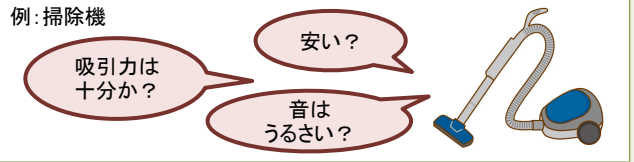
研究の概要

- ◆ 近年、商品レビューを始めとする、消費者によって形成されるビッグデータ(CGM)は膨大なものとなっています。
- ◆ 評判分析(感情分析)は、レビューが示す感情(肯定・否定)を分析することで、**効果的な情報推薦**を可能にします。
- ▶ 本研究では、商品レビューの情報を集約することで、**ユーザの商品選択を支援するシステム**を作成しています。



「評価軸」の自動構築

- ◆ 評判分析に関する多くの関連研究の問題点
レビューを**肯定・否定**という**1次元尺度**でしか分析できない。
- ◆ しかし商品には**複数の評価指標**が存在します。



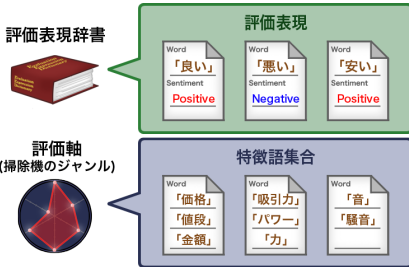
- ◆ 評価指標は「プリンタ」や「掃除機」といった**商品ジャンル毎に異なります**。
- ▶ 新たに**評価軸**を定義、ジャンル毎に学習を行うことで**あらゆる商品ジャンルに特化した評判分析**を可能にします。

評価表現辞書と評価軸

商品ジャンル毎に以下の語彙情報を学習します。

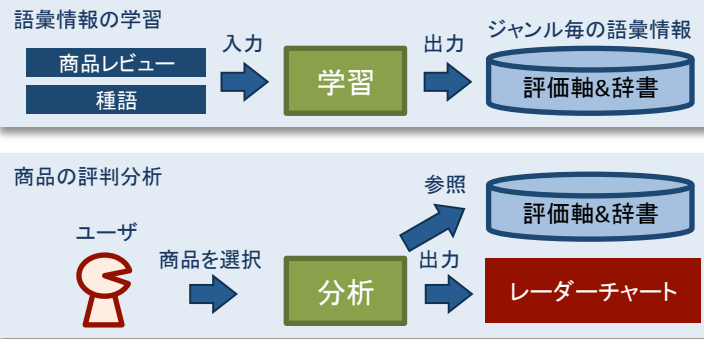
- ◆ **評価表現辞書**
ジャンルにおける肯定語と否定語の語彙集。
- ◆ **評価軸**
ジャンルにおける評価指標、特徴語の集合として定義。

▶ **商品ジャンルに適した評価軸**を自動的に設定でき、レーダーチャートによる可視化が可能に!



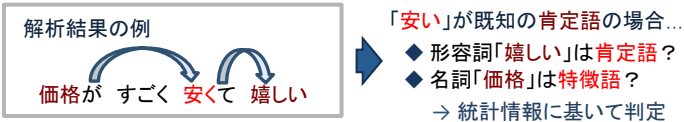
提案システムの流れ

- ◆ **商品レビューの集合と少数の種語**を入力として与えると、**語彙情報を学習**します。
- ◆ ユーザが商品を選択すると、**学習した語彙情報**を利用して商品レビューを分析、**評価の可視化**を行います。



提案手法

- ◆ **種語からの語彙の拡張**
係り受け関係に基づき、既知の表現から新たな評価表現と特徴語を収集します。



評価軸の構築

- ◆ **特徴語のクラスタリング**によって評価軸を構築します。
- ◆ **特徴語間の距離の定義**
特徴語と評価表現の共起回数をもとに、TF-IDF行列を作成します。

	値段	価格	音	吸引力	TF-IDF計算 (正規化)	値段	価格	音	吸引力
安い	10	8	0	0	→	安い	0.88	0.79	0
うるさい	0	0	10	0		うるさい	0	0	1.0
強い	0	0	0	9		強い	0	0	0
満足	3	5	0	5		満足	0.11	0.21	0

▶ **k-meansクラスタリング**により出来たクラスが**評価軸**となります。

成果と今後

- ◆ **現状の成果**
評価軸の構築から、レーダーチャートによる可視化までを一貫して行うシステムは完成しています。「掃除機」ジャンルのレビュー78,504件を対象にした実験では、**評価軸の構築を1分23秒、評価の可視化を3秒**で行うことができました。
また、評判分析の精度評価では**平均80%**の正解率を達成しています。
- ◆ **今後の展望**
提案手法の入力はレビュー集合と少数の種語だけであるため、**あらゆるテキストに対して適用可能**です。今後もCGMは増え続け、効果的な**要約や情報推薦技術**がますます重要になっていきます。この研究を進歩させることで、将来人々の生活を豊かにできる技術に繋がると考えています。

