

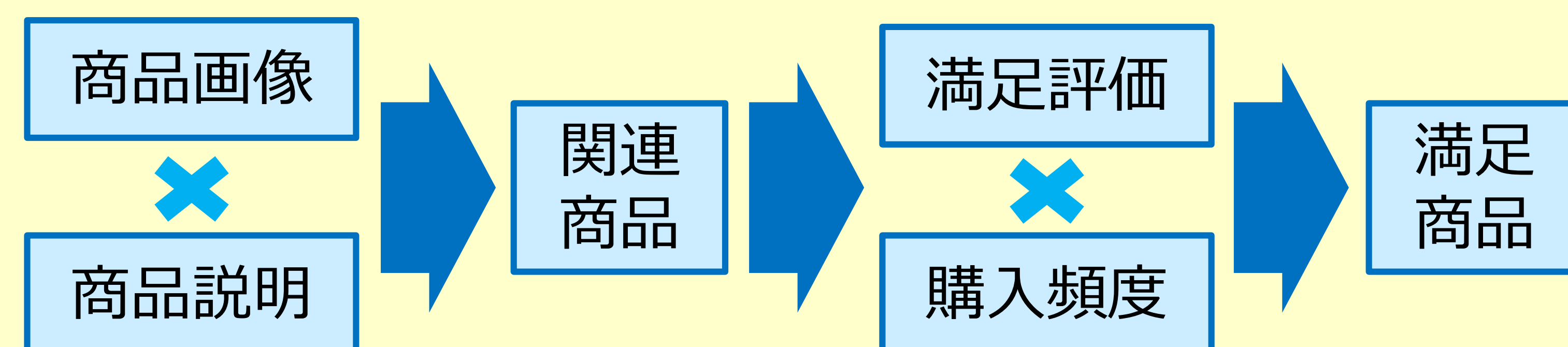
# 画像検索と商品レビュー分析に基づく商品推薦手法の提案

鍾 厚積 王 元元 (山口大学大学院創成科学研究科)

## 商品推薦手法の概要

- ・ Eコマースサイトで商品を購入するユーザが大幅に増加  
例) 楽天、amazon
- ・ 一般検索では商品の具体的な情報が必要  
→ 商品を言葉で表現するのは難しい (例: 服のスタイル)
- ・ 現在の画像検索は商品に対するユーザの不満が多い  
→ 商品画像検索のみでは、関連商品を全て検索できない

目的: 検索された関連商品から満足度が高いものを推薦



## ①商品類似度算出

1. 商品画像のSIFT特徴量を抽出
2. ユーザの入力画像との画像特徴量類似度 (IS) を算出
3. ユーザの入力画像を全ての関連ラベルを追加
4. 商品説明文と画像関連ラベルとのテキスト類似度 (LDS) を算出
5. ISとLDSを正規化し、商品類似度 (S) を算出

$$S = \alpha \times IS + (1 - \alpha) \times LDS$$

( $\alpha = 0.7$ )

## ②商品満足度算出

1. 商品レビューに感情分析\*を行い、レビューごとにポジティブスコア (PS) を算出
2. ユーザのリピート購入行為 (RPW) のレビューに重み付け
3. PSとRPWを用いて、商品満足度 (RS) を算出

$$RS = \frac{\sum_{i=1}^n (PS_i \times RPW_i)}{n}$$

(n: 商品レビュー総数)

\*Google Natural Language API

## ③総合スコア算出

1. 商品類似度 (S) と 商品満足度 (RS) を正規化
2. 推薦商品数 X を決定し、商品類似度上位の 2X の商品を選出  
例) X=10, 上位20件の商品を選出
3. 選出した商品が正規化した商品類似度 (S) と 商品満足度 (RS) を重み付け、総合スコア (OS) を算出

$$OS = \beta \times S + (1 - \beta) \times RS$$

( $\beta = 0.5$ )

## 評価実験: 推薦商品ランキングの検証

- ・ NIIによって提供された「楽天市場データ」\*\*を利用
- ・ 対象商品: 商品画像と説明文を持ち、レビューは10件以上の商品

## 推薦商品ランキングの評価結果

入力画像:

推薦結果上位10件

|              | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    |
|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| ① 類似度ランキング   |       |       |       |       |       |       |       |       |       |       |
| 商品類似度        | 0.702 | 0.699 | 0.697 | 0.695 | 0.686 | 0.685 | 0.679 | 0.673 | 0.671 | 0.667 |
| ② 満足度ランキング   |       |       |       |       |       |       |       |       |       |       |
| 商品満足度        | 1     | 0.978 | 0.925 | 0.899 | 0.849 | 0.836 | 0.827 | 0.811 | 0.742 | 0.717 |
| ③ 制限なし総合     |       |       |       |       |       |       |       |       |       |       |
| 商品満足度        | 0.925 | 1     | 0.978 | 0.899 | 0.849 | 0.811 | 0.827 | 0.836 | 0.742 | 0.698 |
| ④ 類似度TOP20総合 |       |       |       |       |       |       |       |       |       |       |
| 総合スコア        | 1.622 | 1.578 | 1.495 | 1.483 | 1.384 | 1.377 | 1.269 | 1.235 | 1.221 | 1.191 |

入力画像:

推薦結果上位10件

|              | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    |
|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| ① 類似度ランキング   |       |       |       |       |       |       |       |       |       |       |
| 商品類似度        | 0.617 | 0.616 | 0.612 | 0.603 | 0.598 | 0.598 | 0.594 | 0.592 | 0.589 | 0.586 |
| ② 満足度ランキング   |       |       |       |       |       |       |       |       |       |       |
| 商品満足度        | 1     | 0.919 | 0.913 | 0.913 | 0.826 | 0.826 | 0.804 | 0.696 | 0.696 | 0.645 |
| ③ 制限なし総合     |       |       |       |       |       |       |       |       |       |       |
| 商品満足度        | 1     | 0.919 | 0.913 | 0.913 | 0.826 | 0.826 | 0.804 | 0.696 | 0.696 | 0.645 |
| ④ 類似度TOP20総合 |       |       |       |       |       |       |       |       |       |       |
| 総合スコア        | 1.538 | 1.481 | 1.419 | 1.402 | 1.393 | 1.358 | 1.266 | 1.212 | 1.211 | 1.207 |

| ランキング       | MRR   | MAP   | NDCG@10 |
|-------------|-------|-------|---------|
| ①類似度        | 0.529 | 0.516 | 0.625   |
| ②満足度        | 0.556 | 0.668 | 0.679   |
| ③制限なし総合     | 0.561 | 0.626 | 0.745   |
| ④類似度TOP20総合 | 0.594 | 0.757 | 0.936   |

- ・ ④類似度TOP20総合ランキングは最も良い結果
- ・ 10名の被験者に総合ランキングに対する評価
- ・ ③制限なし総合 : ④類似度TOP20総合 = 1 : 9

## 今後の課題

- ・ 画像特徴量類似度算出の改善
- ・ 画像関連ラベル抽出手法の改善
- ・ ユーザのリピート購入行為検出方法の改善

謝辞: 本研究では、楽天株式会社が国立情報学研究所の協力により研究目的で提供している「楽天公開データ」を利用した。ここに記して謝意を表す。

\*\*楽天株式会社. 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>, 2014年.