

# P09 楽天レビューにおけるアイテムレビューとショップレビューの混同に関する実験的分析

土屋佳未, 吉岡真治 (北海道大学)

## 1. 研究背景・目的

### ■ 楽天レビューの仕様

アイテムレビューとショップレビューは分けて書く

お得です。 **アイテムレビュー** アイテムレビュー記載欄に書かれたレビューの一例  
 服5点、雑貨1点のときに購入しました。

雑貨を含め捨てアイテムがなく、大変気に入りました! **ショップレビュー**  
 金額訂正連絡の時に、ミスがありました。問い合わせるとすぐに対応してくださいました。  
 また機会があれば、購入したいです^^

### ■ 目的

アイテムレビューとショップレビューの分類をする仕組の提案

## 2. データの扱い方

**アイテムレビュー** → 買った商品に関するレビュー  
 使いやすさ、デザインなど

**ショップレビュー** → 買った店舗に関するレビュー  
 配送、スタッフの対応など

## 3. 方法

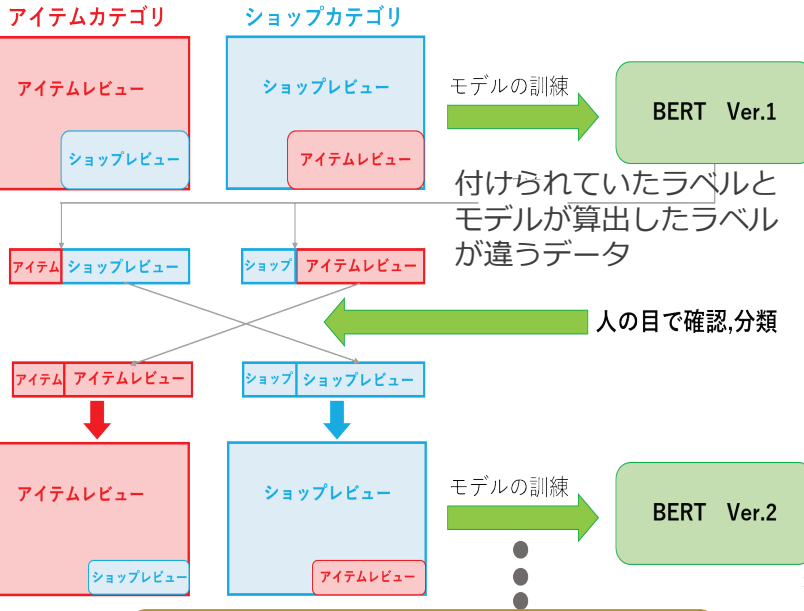
### 機械学習

自然言語処理モデルBERTを使用

入力: レビューに書かれた文

出力: アイテムかショップかのラベル

### ■ 訓練データの作り方の提案



データセットのクリーニングとBERTの訓練を逐次的に行う

## 4. 実験・結果

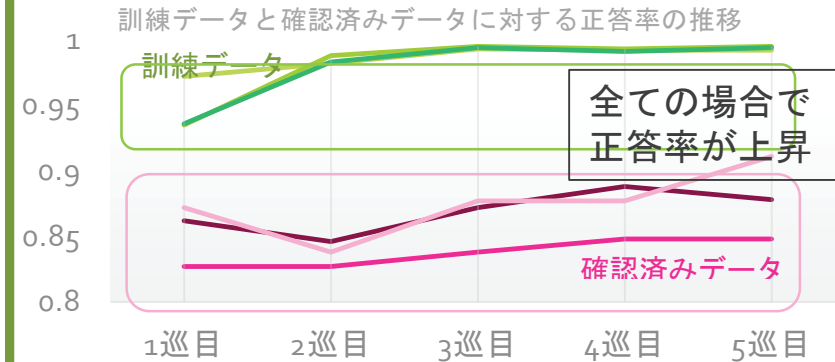
### ■ 実験1: 事前調査

ランダムサンプル200件中どのくらいノイズ(書かれたカテゴリが違う文)が入っているか調査

実験	ノイズ数(個)	割合(%)
1	27	13.5
2	25	12.5
3	29	14.5
平均	27	13.5

### ■ 実験2: 分類モデルの精度の推移

訓練データとすでにクリーニングしてある確認済みデータに対してBERTの精度を検証



### ■ 実験3: 最終実験

実際のデータのノイズに対する検出精度の検証

実験	実際のノイズ数(個)	BERTが見つけたノイズ数(個)	割合(%)
1	27	23	85.2
2	25	20	80.0
3	29	27	93.1
平均	27	23	86.1

## 5. まとめ

1. アイテムレビューとショップレビューの混同が多く発生していることを確認
2. 二つのレビューの分類システムの構築とトレーニングデータの修正を繰り返す方法を提案し、その有効性を確認