

Sentence BERTによる出品者特性を考慮した売却時間予測モデルの提案

s15 金田翔平, 佐藤優月, 鞆津悠二, 馬越紬未, 久保幸平, 野口哲汰, 宿久洋 (同志社大学)

背景

商品説明文に商品画像のような視覚的な情報では説明できない特徴を記載することで、EC サイト利用者の購買意欲の向上につながる[1]

メルカリの商品説明欄には、様々な情報を載せることができ
同じ商品でも多様な書き方がある

商品説明A (商品内容重視)

イタリア製です。
中古品です。
茶色の27cmです。



商品説明B (対応重視)

値下げ可能
たのメル便での発送
送料負担

どんな書き方、情報が購入者に効果的に働いているか分からない

目的

- ①出品者特性を反映させた特徴量を基に、**出品時に売却時間を予測するモデルを構築**
- ②各特徴量の**影響度を解析**

▶ **購入を促す商品説明文の特徴を明らかに!**

使用データ

メルカリデータセット

- ・商品のステータスが**sold_out**
- ・商品カテゴリが**バッグ**

計**139370**件

モデル概要

特徴量選定

目的変数：**出品時から商品売却までの時間**

説明変数：**出品者特性**を反映させた以下の特徴量



・画面上の行数・密度



▶ スマホ画面上での改行文字数(24字)を基に
画面上での行数及び密度を計算

行数：自分の意思で改行した場合と仕様上の場合を考慮する必要

行数 = **改行の個数 + 一行が24字以上の文数**

密度 = **総文字数 ÷ (行数 × 24)**

・文末表現



敬語

新品未使用です

体言止め

新品未使用

その他

新品未使用だ

※文末表現で出品者への印象が変化する可能性

▶ **形態素解析を用いて文末表現を判別**

・出品者が業者か個人か



▶ **業者か個人かでラベル付け**

基準：商品説明文の内容の重複の有無

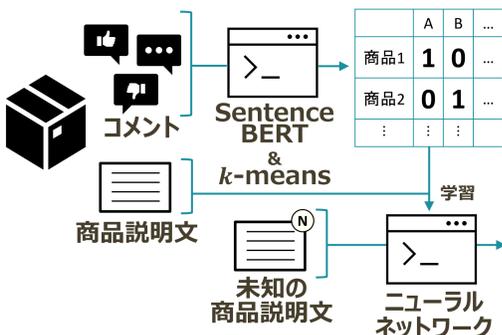
※業者は同じ商品に対して、コピー&ペーストで説明欄に記載する

その他の変数

- ・商品の値段
- ・文字数
- ・コメント数
- ・送料負担の有無
- ・発送方法
- ・発送期間
- ・土日祝に出品されたか

※土日祝は平日に比べて
利用者が多いことが想定される

・商品に対するコメント (質問・交渉)



未知の
コメントの
要素を推測

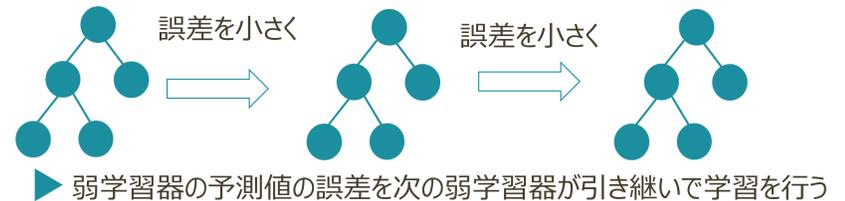
足りない
と推測される
要素を表す
ラベル

※コメントは商品説明文の不足により発生すると仮定

ラベル概要

ラベル	不足内容
0	商品状態
1	即購入可か否か
2	商品の装飾
3	画像の鮮明度
4	サイズ/色味

勾配ブースティング



モデルの選択

使用モデル：「**XGboost**」

勾配ブースティングアルゴリズムを用いた機械学習法

▶ **XGboostに上記の特徴量を適用**

結果・考察

SHAP値を用いた解釈

機械学習モデルの各特徴量の影響度を測るための手法

売却時間に影響を与える特徴量(上位)

- **商品状態に関する説明の充実度合**
 - ・ **“商品状態”に関する説明が不足していると長くなる傾向**
- **文字数**
 - ・ **文字数が短いと長くなる傾向**
- **画面上の密度**
 - ・ **密度が低いと長くなる傾向**
- **体言止め**
 - ・ **体言止めを用いると長くなる傾向**

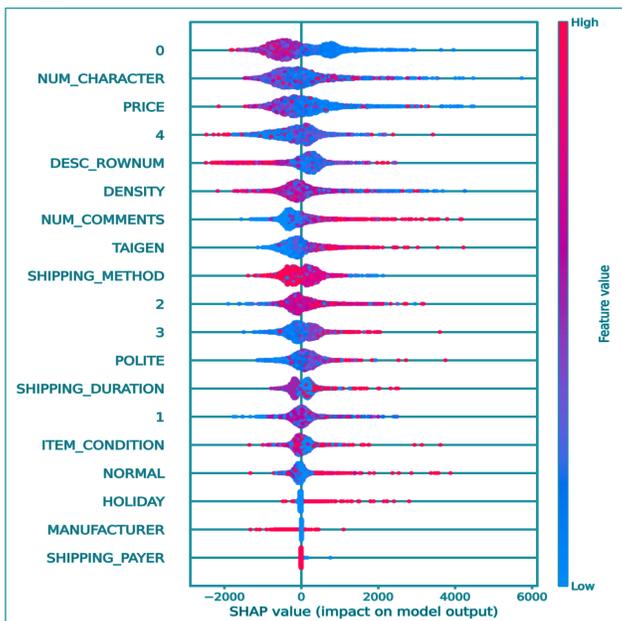
特に上記の項目を重視した商品説明欄づくりが必要

今後の展望

- ・ **モデル予測精度の向上**
より説明力の高い特徴量の選定
ex) 商品画像
- ・ **モデルを用いた施策の提案**
ex) 商品説明欄の採点機能の導入

謝辞

本分析では、株式会社メルカリの協力のもと、国立情報学研究所より「メルカリデータセット」を利用させていただきました。ここに記して謝意を表します。



SHAPによる各特徴量の影響度

[1]鈴木凜次, 福本健二, 寺田浩之, 馬頭正文, 灘本明代, “商品説明文自動生成のための深層学習モデル比較検討”, 第13回データ工学と情報マネジメントに関するフォーラム (DEIM2021), E25-2,7pages, 2021.