

局所画像特徴量を考慮した中古ジュエリーの商品売値推定モデル

津村 大悟¹, 大竹 恒平², 生田目 崇³

キーワード: openCV, SIFT 特徴量, ニューラルネットワーク, Entity Embedding, XGboost

1. はじめに

現在、リユース商品を購入するユーザが増加し、リユース市場は拡大している。リユース事業を展開する企業の多くは、ユーザから提供される中古アイテムに対し、出品者が買取価格と出品価格を決定し、再販売を行う。この時、出品価格は出品者の経験に基づいて意思決定されていることも多く、必ずしも最適な価格決定がされているわけではない。提示する商品価格と消費者の期待する価格との乖離が少なければ、商品の流動性が上がり、結果として売れ行きがよくなるなどの好循環が期待できる。以上より、適正な出品価格を決定することはビジネス上大変重要な意思決定になってきている。本分析では、中古ジュエリーを取り上げ、EC においては特に重要と思われる画像情報を考慮した価格推定を行う。

2. データと方法

本研究では主にリユースジュエリーの買取・販売を行っている企業から提供いただいたデータを使用する。また、分析内では、2019年9月1日～2020年8月27日の期間内に販売されたジュエリー4142件に関する購買データ、商品属性データ、そして商品の画像データを用いる。本研究では、ニューラルネットワークと XGboost モデル [1] を用いて価格予測を行い、2つのモデルでの精度比較を行う。ニューラルネットワークでは、Entity Embedding を用いることで、カテゴリカル変数の表現を豊にしている。また、画像データに関して、openCV を用いて SIFT 特徴記述子を抽出し、輝度に関する特徴量を分析に用いた。

3. 結果と考察

2つの手法による予測結果を図1, 図2に示す。

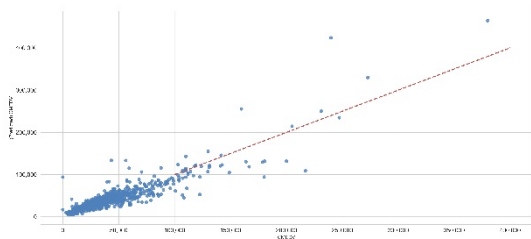


図 1: Entity Embedding モデルによる価格予測結果

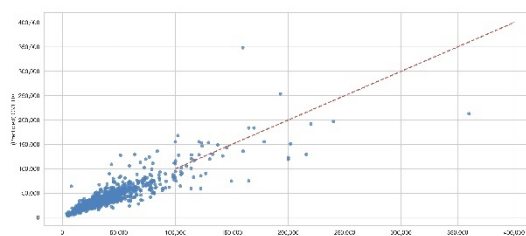


図 2: XGboost モデル 価格予測結果

また、精度評価は以下のような結果となった。

表 1 精度評価

| 評価指標 | Entity Embedding | XGboost |
|---------|------------------|---------|
| 平均絶対誤差率 | 22.49 | 23.57 |

以上より、Entity Embedding モデルを用いた予測の方が、XGboost を用いた予測よりもやや精度の高い予測ができていることがわかる。しかし、価格が高い中古ジュエリーに関して、うまく売値価格を予測することができなかった。要因として、定価よりも売値の方が高い中古ジュエリーが一部存在することが考えられる。

また、XGboost を用いて、予測に影響を与えていた変数を調べた。その結果、商品の定価価格の他に、画像特徴量が予測に大きな影響を与えていたことが明らかになった。

4. おわりに

本分析では、Entity Embedding モデルと XGboost モデルを用いて中古ジュエリーの価格推定を行った。

ニューラルネットワークを用いた予測では、Entity Embedding を用いたことで、カテゴリカル変数をうまく予測に用いたことで、XGboost よりも高い精度で予測することができた。

また、画像特徴量が予測に効いていたことから、画像特徴量の有用性が明らかになった。今後の課題として、特徴量選択について改善を行う必要があると考える。今回は 128 次元の特徴ベクトルについて、最大値を選択し、説明変数として利用したので、より適正な特徴量選択を行うことで、より予測の精度をあげることができると考える。

参考文献

[1] T. Chen, C. Guestrin, "XGBoost: A Scalable Tree Boosting

1 中央大学大学院
2 東海大学
3 中央大学

System”, KDD '16: Proceedings of the 22nd ACM SIGKDD
International Conference on Knowledge Discovery and Data
Mining, pp. 785–794 (2016)