

S19 機械学習を用いた製品画像の取引成功率における影響

芦邊快斗, 佐藤龍ノ介, 長谷俊吾, 伊集院大将 (千葉工業大学)

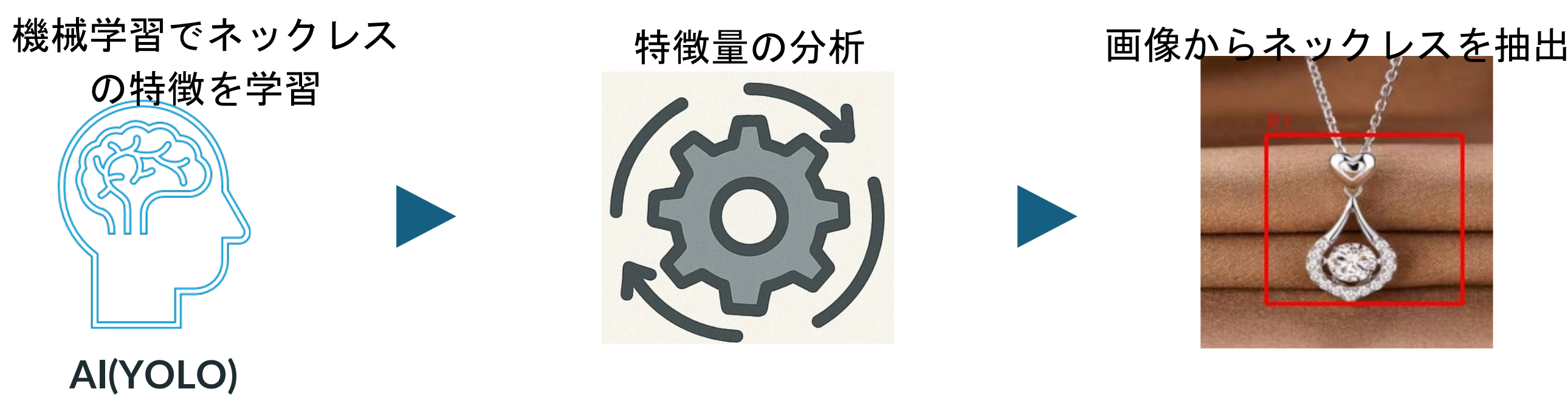
研究背景

ECサイトの画像におけるいい写真というのは撮影者の主観に依存し、定量的な基準が存在しなかった。

研究目的

様々な角度から機械学習を用いて商品画像を分析し、取引成功率を調査する。

注目度の抽出過程



データの前処理

- ・使用データ：メルカリデータセット
- ・対象商品：ネックレス
- ・ノイズ除去（価格の制御）



価格の歪みが大きかったため、高価格10%と低価格10%を除外。

価格帯（550円～7610円）に絞り込み、対象商品118件を分析対象とした。

仮説

- ・視覚的特徴が購買に影響を与えるのか
- ・売れた商品の構図は販売スピードに影響はあるか



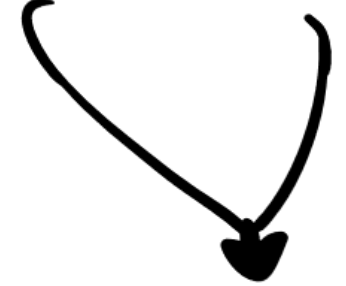
分析1：画像特徴を用いた取引成立予測モデル

目的

商品画像から売れやすさを推定できるモデルの構築と、画像特徴が取引成立にどの程度影響するかの検証

手法

1. 画像特徴量抽出
2. 次元削減
3. モデル構築



結果

表1：画像特徴を用いたモデル比較結果

Model	Accuracy	ROC-AUC	MSE	R ²
Logistic_img	0.688	0.609	-	-
Linear_img	0.562	0.512	0.241	-0.192
RandomForest_img	0.719	0.449	0.218	-0.081

① Accuracy (正解率)

画像だけでは安定した分類性能は得られなかった。

② ROC-AUC (判別性能)

売れやすさを判別する力は弱かった。

③ MSE (誤差)

予測誤差が大きく、回帰的な予測は難しかった。

④ R² (説明力)

画像特徴だけでは売れやすさを説明できなかった。

分析2：視覚的特徴の可視化を用いた売れ行き要因の分析

目的

商品画像の構図や画質指標（中央集中度、コントラスト、背景占有率）が販売スピードに与える影響を検証する

手法

1. 特徴量抽出
2. 画像指標算出
3. AIの注目度と中心のズレの指標をしめす 中央集中度 (CentralityScore)

$$distance = scaled - 0.5$$

$$total_distance = \sqrt{(distance_x)^2 + (distance_y)^2}$$

$$CentralityScore = 1.0 - total_distance$$

結果

1. コントラスト平均
24.25 (売れた) 23.56 (売れていない)
2. 背景占有率平均
69% (売れた) 68% (売れていない)
3. 中央集中度平均
0.52 (売れた) 0.51 (売れていない)

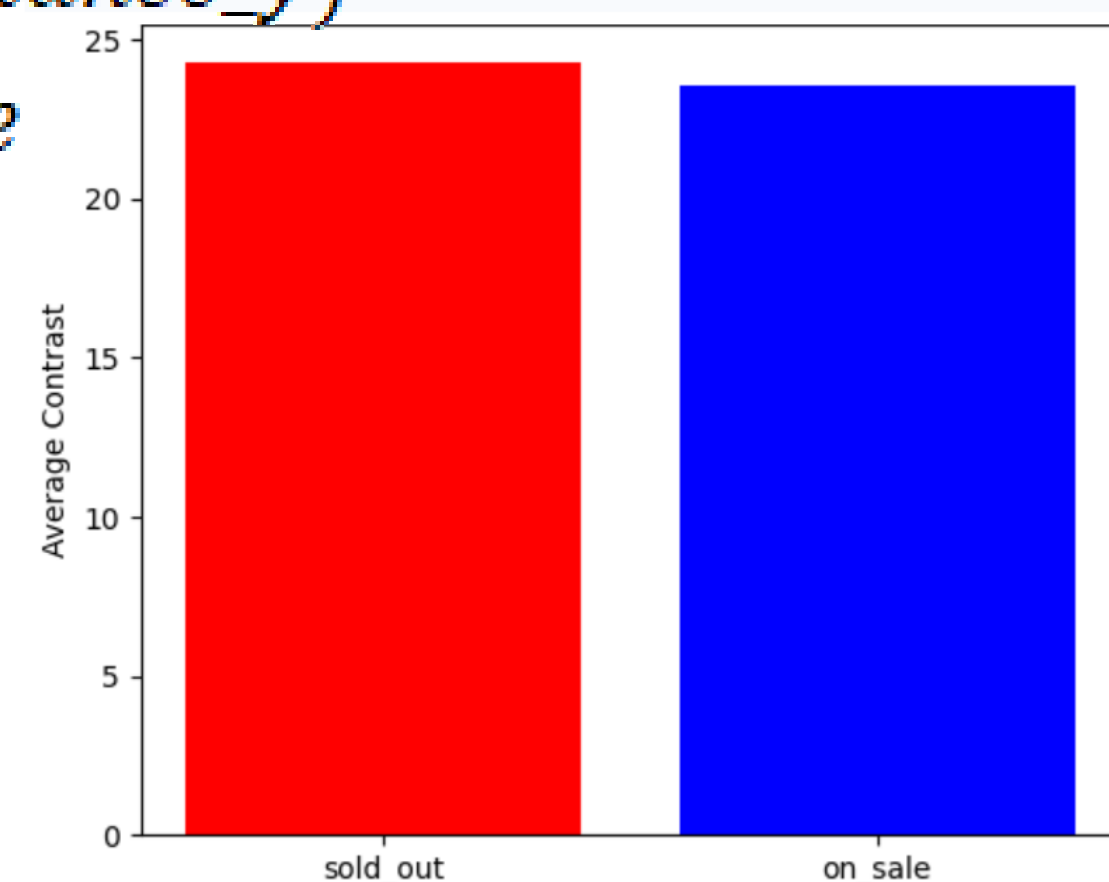


図1.コントラスト平均

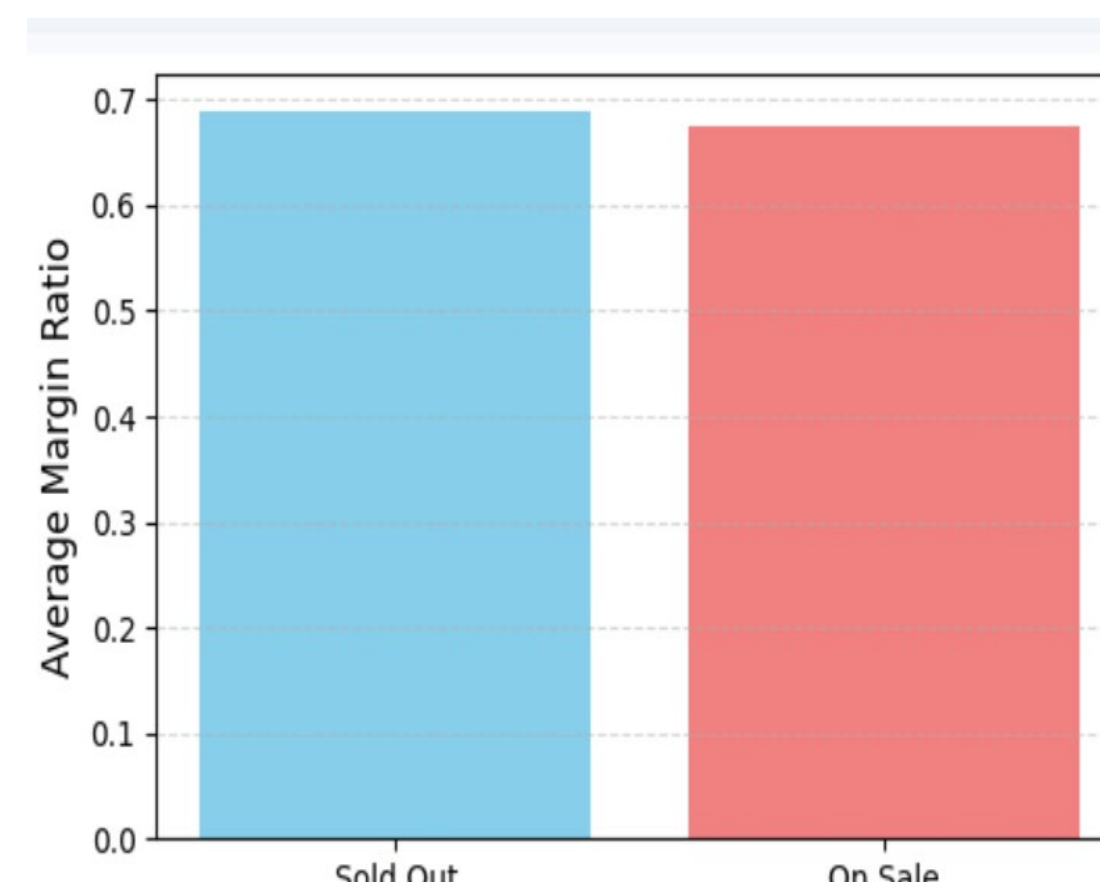


図2.背景占有率平均

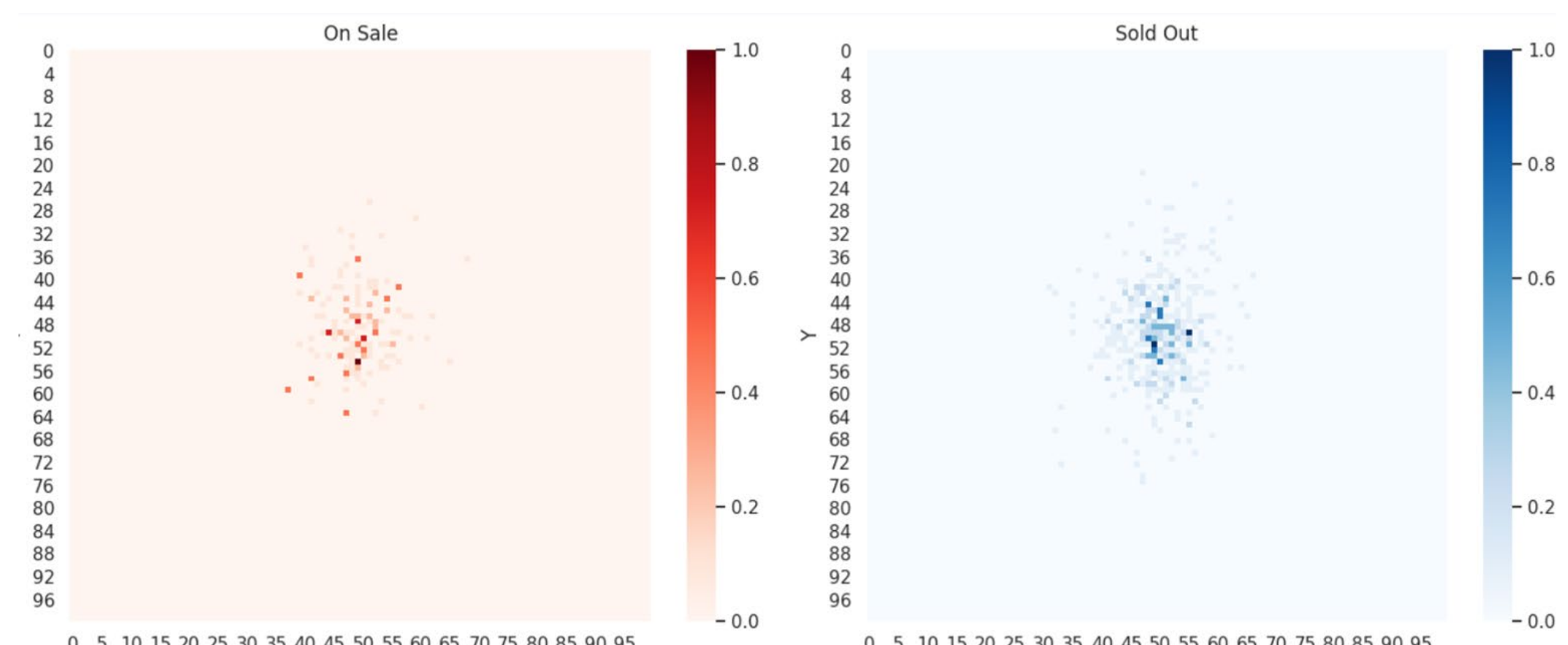


図3.中央集中度

- 判明したこと：背景占有率・コントラスト・中央集中度は売れやすくなるため要素ではなかった

分析3：AIによる構図の定量化データを用いた構図と販売スピードによる分析

目的

売れている画像構図と販売スピードの関係性が加速要因または減速要因として働くのかを検証

手法

- ・特徴量抽出
- ・統計的検定 (t検定)

結果

- 分析A：N=55 相関係数 $r = -0.0654$ 、 p 値 = 0.6350
判明したこと：統計的に有意な差はなかった。
分析B：Fast群(<=6日, N=40) 平均スコア：0.4750
Slow群(>=60日, N=40) 平均スコア：0.3903
t検定の結果： p 値 = 0.0262
判明したこと： p 値が0.05を下回ったことから、0.0847の差は「偶然ではなく、有意な差」である。

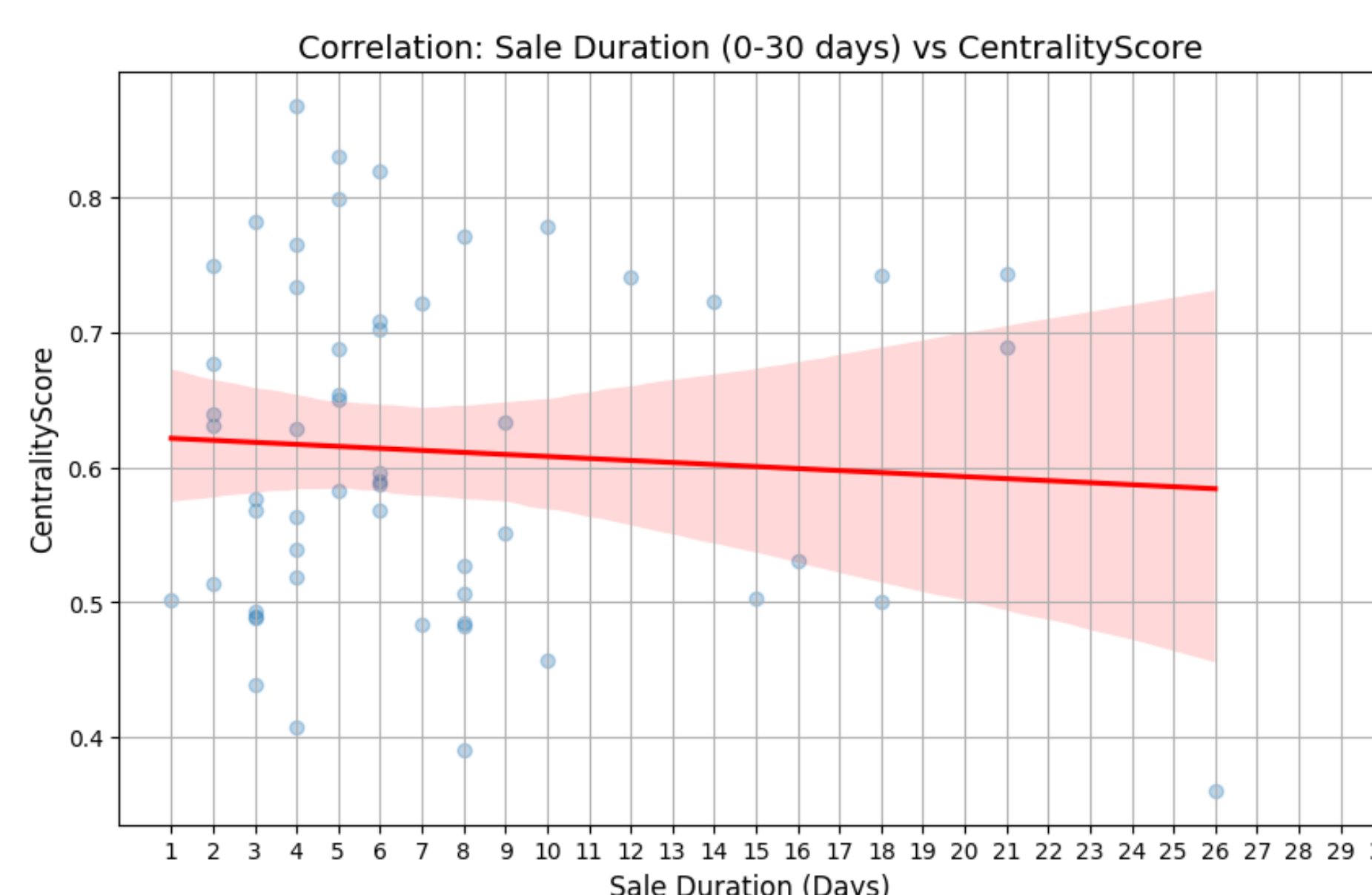


図4 中央集中度と日数の関係(30日間)の相関図

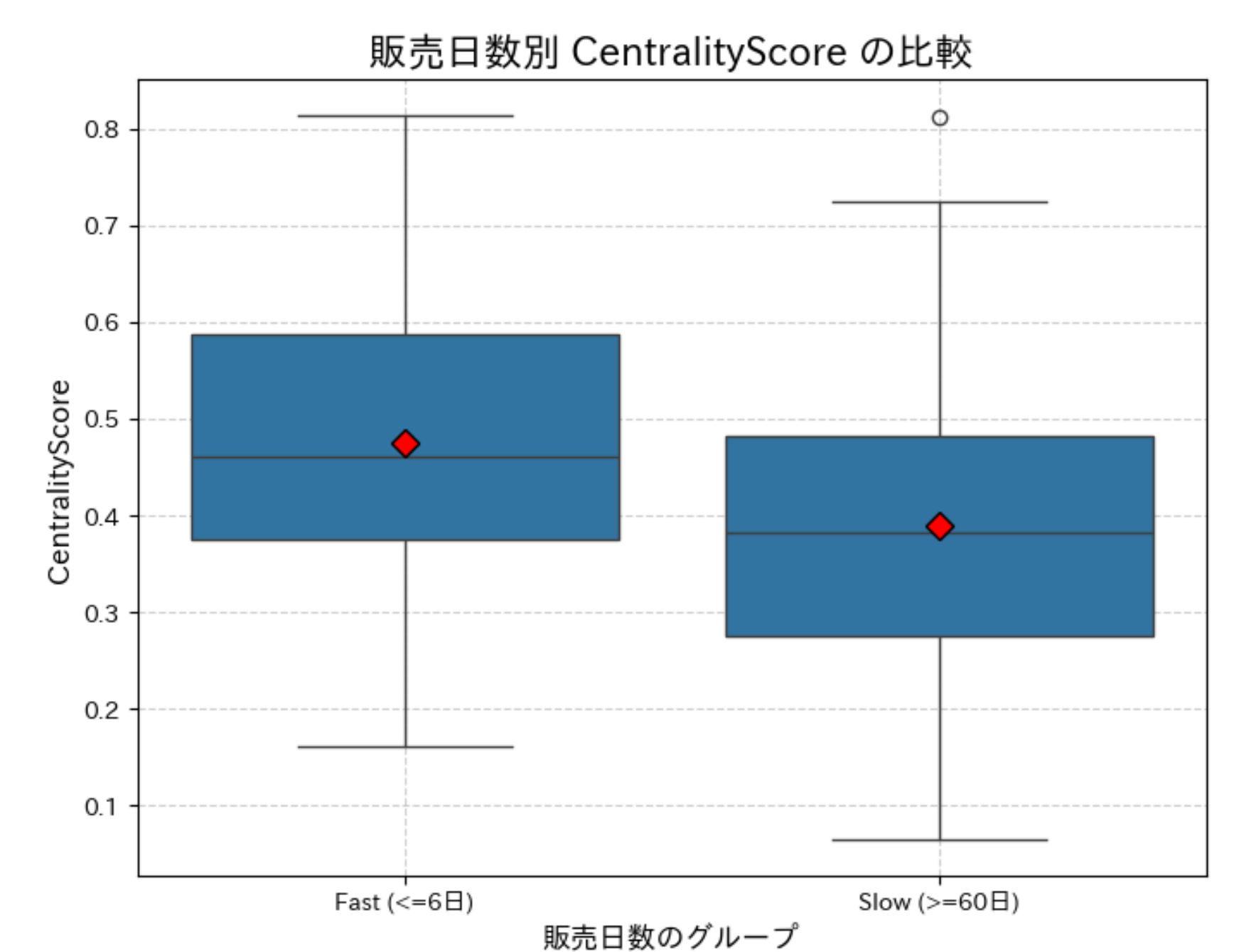


図5 販売スピードごとに分けた中央集中度の箱ひげ図

今後の展望

1. 価格要因を分離した重回帰分析による仮説の厳密検証 今回は価格帯を絞り込むことで価格の影響を低減したが、今後は価格自体を変数に加えた重回帰分析を行う。これにより、価格の影響を統計的に統制した上で、構図スコアが「単独で」販売を長期化させるブレーキとして機能しているかを厳密に検証する。

2. 分析指標の拡大と実用化 (ガイドライン開発) 本研究で用いた「中央集中度」に加え、「明るさ」「色彩」など他の視覚的特徴量や、商品説明テキスト等の情報も組み合わせた分析へ拡大する。将来的には、売れ行きに影響する要因をスコア化し、出品者が改善できるガイドラインとして実用化することを目指す。