Received: 10 August 2024 | Revised: 8 February 2025 | Accepted: 14 March 2025 | Published online: 26 March 2025

RESEARCH ARTICLE

Journal of Data Science and Intelligent Systems
2025, Vol. 00(00) 1–9
DOI: 10.47852/bonviewJDSIS52024073

BON VIEW PUBLISHING

Identifying Purchasing Factors in Online Flea Markets Considering Thumbnail Images

Emi Iwanade[1], Yoshihisa Shinozawa[2] and Kohei Otake[3,*]

[1]*Graduate School of Information and Telecommunication Engineering, Tokai University, Japan*

[2]*Faculty of Science Technology, Keio University, Japan*

[3]*Faculty of Economics, Sophia University, Japan*

Abstract: In recent years, the market for online flea markets, which are consumer-to-consumer (C2C) services where goods are bought and sold among users, has been expanding. In such services, sellers (individuals that offer goods or services for sale) create product details, including price, condition, shipping method, and images, when listing their items for sale. We consider the product image to be the first thing users see when selecting a product as a thumbnail, significantly impacting their purchase decisions. In this study, we proposed a discriminant model of purchase decisions for online flea market data to clarify the factors influencing these decisions based on product details. Specifically, we used metadata such as price and delivery method, along with image labels for product thumbnails, as features. We created and compared models with three patterns: metadata only, image labels only, and a combination of metadata and image labels. We created four types of models in this study: logistic regression, decision tree, gradient boosting, and random forest. We selected the models based on their accuracy evaluations. Our analysis revealed that the model using both metadata and image labels as features, combined with the gradient boosting method, had the highest accuracy. The partial dependence plots of the selected models highlighted the features important for users' purchase decisions.

Keywords: online flea markets, images, machine learning, gradient boosting, consumer behavior

1. Introduction

In recent years, the market size of online flea markets, which are consumer-to-consumer (C2C) services for buying and selling goods among users, has been expanding. According to a survey on e-commerce conducted by the Daiwa Institute of Research [1], the market size increased from JYP 1.147 trillion to 1.243 trillion from 2020 to 2021, representing an increase of about 22.5% from the previous year.

Generally, in an online flea market, sellers register photos, descriptions, and prices of their items for sale. The buyer then selects a product they like and makes payment through the platform. Next, the seller ships the item to the buyer and shares the tracking number and other information. Finally, the buyer receives the item and the transaction is complete. In some cases, transactions in online flea markets are anonymous, making them secure in terms of personal information. They also have the advantage of not requiring a storefront and of making it easier to find unique products.

In this study, we focus on online flea markets, which are in growing demand. We create a discrimination model for online flea market data and identify the factors that affect users' purchase decisions by interpreting the model's feature values. Specifically, we use metadata such as price, delivery methods, and image labels for thumbnail images of products as features. We create and compare models using three patterns: metadata only, image labels only, and metadata and image labels. Additionally, we created four types of models: logistic regression, decision tree, gradient boosting, and random forest, and selected one based on an accuracy evaluation.

Through this study, we aim to analyze critical features, including thumbnail images, in online flea markets to uncover what sellers should consider when listing their products to facilitate quicker and easier sales. The key contributions of this study are as follows:

1) Identification of Features Influencing Purchases:

We identify features that significantly impact purchasing decisions in online flea markets. By incorporating these features into their listings, sellers can enhance the likelihood of faster and more efficient sales. This enables sellers to generate income more quickly and secure funds for subsequent listings immediately. Additionally, quicker sales reduce the burden of managing unsold items, allowing sellers to operate more efficiently and begin new sales cycles sooner.

2) Insights into Thumbnail Image Features:

Beyond metadata analysis, this study focuses on the attributes contained within thumbnail images, which are the first point of contact for consumers when selecting products, to identify factors that enhance the likelihood of purchase. By designing thumbnails with these findings in mind, sellers can provide consumers with more attractive and relevant visual cues. This improves the

*Corresponding author: Kohei Otake, Faculty of Economics, Sophia University, Japan. Email: k-otake@sophia.ac.jp

decision-making process for consumers, enhancing their overall satisfaction with the purchasing experience.

3) Activation of Online Flea Markets:

By addressing the needs of both sellers and consumers, as outlined above, this study is expected to stimulate the activity of online flea markets. This, in turn, contributes to the promotion of sustainable economic activities by creating a more dynamic and efficient marketplace.

4) Advancing Image Feature Analysis Specialized for Online Flea Markets:

In this study, we analyze image features by categorizing them into multiple elements unique to online flea markets. Specific examples include edited images with overlaid text on clothing, clothing displayed on hangers, and items folded and placed on the floor. Methods such as deep learning-based optical character recognition (OCR) could be employed to extract these complex features. However, capturing all unique elements would require a large, labeled dataset and significant computational resources. To address this, we utilize a manually labeled dataset to analyze and identify factors deemed critical for influencing purchase decisions. The factors identified in this study are expected to serve as foundational knowledge for future research involving large-scale datasets aimed at image feature extraction. These insights will provide a basis for developing advanced analytical techniques tailored to the characteristics of online flea markets.

In Section 2, we review existing research on online flea markets and machine learning and image analysis. We summarize the characteristics of online flea markets, recent work on predictive analysis using machine learning methods and thumbnail images, and we position how the research brings new insights. In Section 3 we describe the details of the datasets and labeling items used. In Section 4, we train our four models: logistic regression, decision tree, gradient boosting, and random forest, and describe our comparative results. In Section 5, we identify the factors that contribute to purchasing using the model with the highest accuracy. Finally, in Section 6, we summarize our results and discuss the significance of this study and future research directions.

2. Literature Review

With the expansion of online flea markets, research focusing on the characteristics of these platforms has been active. Hellianto et al. [2] conducted a survey of users of Tokopedia and Bukalapak, which are C2C e-commerce websites, investigating what would create a better user experience. The results showed that it is important for online flea market websites to be simple and easy to use. İbrahim et al. [3] researched users' motivations for purchasing used goods online and found that word-of-mouth about products on the Internet is important for fostering sustainable consumption awareness and promoting the purchase of used products. Moriuchi et al. [4] focused on C2C e-commerce and conducted an analysis using questionnaires on satisfaction and trust in transactions, in addition to emotional and functional values when purchasing products. The results revealed that communication between sellers and buyers is important in C2C.

Additionally, Okamoto et al. [5] explored the perspectives users consider important when using online flea markets, while Jenniina et al. [6] investigated how users judge the value of goods on these platforms. Their findings indicate that users tend to prioritize the management system of personal information, followed by the commission fee. Factors such as product price and the elapsed time since the product was sold also influence users' perceptions of product value, with the impact of elapsed time varying depending on the product brand. These studies suggest that there are specific user characteristics unique to online flea market platforms where used goods are bought and sold.

One of the characteristics of online flea markets is that sellers (individuals that offer goods or services for sale) set product information such as price, condition, shipping method, and product images when listing their products. Among these, product images have been shown to significantly impact purchase decisions [7]. Zhong et al. [8] created a product recommendation model using image information in e-commerce, demonstrating that a model combining product review text scores with product image information provides better product recommendations to users. Previous research [7, 8] has shown the usefulness of image information in making purchase predictions. However, these studies often use features obtained from trained convolutional neural network models for image information, considering multiple features simultaneously. This approach does not account for important viewpoints in planning marketing measures.

In this study, we create features that we expect users to consider when making purchase decisions for product images and attempt to create a model that enables users to evaluate these features independently. Predictive analysis has been conducted in various fields using machine learning techniques. Valavan et al. [9] used decision tree, random forest, logistic regression, and gradient boosting to predict credit card defaulters. The results showed that the model using gradient boosting was the best in several evaluation indices. Wu et al. [10] proposed the best combination of feature extractors and classifiers for purchase prediction by combining multiple machine learning methods using customer purchase history data. The results showed that the combination of gradient boosting and ANN can output more accurate prediction results. Furthermore, Ghosh et al. [11] used a random forest model to clarify what is important as a purchasing factor based on customer purchase data and advertising log data. As a result, they found that the number of clicks and the order of advertisements were the most important factors, and services used in the past were not so important. The results show that machine learning has been used to make various predictions about consumer behavior, and that methods such as random forests and gradient boosting in particular are frequently utilized.

Researchers study machine learning methods on a daily basis, as shown in previous studies [9–11], and the effectiveness of these methods depends on the dataset used. Therefore, it is necessary to try multiple methods, evaluate their accuracy, and select the most fitting model for consideration. For example, Zhong et al. [12] compare four methods, extreme gradient boosting, random forest, artificial neural network, and logistic regression, for the purpose of predicting outcomes in critically ill patients after open-heart surgery. In this study, the easy interpretability of the results is essential due to the characteristics of this research on the understanding of consumer behavior in the marketing domain. Therefore, based on previous research, we compare four models: random forest, gradient boosting, logistic regression, and decision tree, which can be interpreted in a rule-based interpretation.

Online flea markets typically use images captured by sellers, or quoted from official product websites. In the case of an image taken by a seller, we can read various elements from a single image, such as the presence or absence of a product logo, the presence or absence of packaging, whether the image is hung on a wall or placed on the floor, and whether text has been added to the image through editing.

In this study, we focused on the thumbnail images, which are the first point of contact for consumers when selecting products. We consider thumbnail images to have a significant impact on users' purchase decisions because they are the first images that users see on a product page.

Sato et al. [13] conducted a YouTube study using thumbnail images. They first categorized video thumbnail images according to the genre. Then, they prepared samples of thumbnail images with multiple patterns, including the size of the text and whether to include people, and revealed what viewers considered important when watching a video. The results indicate that the saturation of the thumbnail images affected the selection of videos to be viewed. Zhang et al. [14] conducted a similar study using YouTube thumbnail images, focusing on the interests of different countries. The results revealed differences in video categories of interest between Asians and other countries. In addition, there are many other studies focusing on thumbnail images, such as Miyamoto's [15] research that attempted to classify thumbnail images based on the number of accesses using machine learning and found that there is a certain trend between thumbnail images and the number of accesses. Kozu et al. [16] experimentally evaluated the difference in sensitivity to thumbnail images in online video sharing services for tourism. These studies show that thumbnail images are the first thing that users see when they select things and are an important factor in comparison and examination.

3. Dataset

In this study, we use the "Mercari Dataset" provided by the Informatics Research Data Repository of the National Institute of Informatics, which is related to the online flea market application Mercari [17]. The dataset contains 17 product metadata items such as listing status and price, product comment data, and product image data. Metadata and imaging data were used.

3.1. Extraction of target products

Online flea markets offer several product categories. In particular, clothing is often photographed with mannequins or folded on the floor, and photos taken by sellers tend to show differences among multiple product categories. In this study, we analyze online flea markets by considering these unique features. Methods such as deep learning-based optical character recognition (OCR) could be employed to extract these complex features. However, there are several factors to consider when dealing with thumbnail images used in online flea markets. For instance, these images are freely taken by consumers, resulting in data that often contains noise due to inconsistencies in format. Additionally, there are numerous potential factors that may influence outcomes, making it challenging to identify all relevant features comprehensively. Additionally, the predicted labels tend to include more errors compared to labels classified manually. For these reasons, this study utilizes manually labeled data.

As an initial step to address this challenge, this study involves creating a manually labeled dataset for images and developing a model aimed at identifying critical features that influence purchase decisions. By utilizing the critical features identified in this study, the next step of extracting features using OCR can be conducted more efficiently and with prioritized focus, enabling a more effective exploration of relevant features.

In this study, we used the fashion category as the data category and selected the tops of men and women, which have the largest number of transactions. Among the items on display, we targeted the two brands with the highest number of transactions: fast fashion and its sister brand. In addition, we selected items priced between JPY 300 and JPY 10,000, considering outliers.

In addition, among the target commodity transactions, we reduce the data by random sampling from a data- processing perspective. In this study, we extracted 3,000 records, 1,500 of which were in the sold state and 1,500 in the on-sale state, to classify transactions based on binary values according to whether they were purchased.

3.2. Creating datasets

In this study, we used the metadata and image labels accompanying all 3,000 records as product features. Specifically, we used the following metadata: listing status, price, number of likes, shipping method, shipping origin region, number of days until shipping, and product condition. For the listing status, which we used as the objective variable, we set 1 for sold and 0 for currently listed; for the other metadata, we performed one-hot encoding on all features except price and number of likes.

Table 1 presents an overview of the tagged image labels. Three university students with experience in using Mercari, including the author, visually checked the thumbnail images of the products and labeled them as extractable from the 15 fashion category

Table 1
Types of image labels, their labeling methods, and the number of label occurrences

	Image label	Criteria	Number of labels
Text Object on Image	Size	The image includes the clothing size	254
	Brand Name	The image includes the brand name	142
	Delivering Method	The image includes about delivering method	5
	Discount	The image includes about discount	21
	Brand's Logo	The image includes the brand's logo	11
Situations when Taking Pictures	Floor stand	Clothing is photographed on the floor	1,519
	Hanger Hanging	Clothing is photographed on a hanger	981
	Mannequin (Person) Wearing	Clothing is photographed on a mannequin or person	42
	Folded	Clothing is photographed in a folded	139
	Packaged	Clothing is photographed in a packaged	174
	Price Tag	Clothing is photographed with a price tag	283
Type of image	Official Website Image	Images from websites are used as thumbnail images	400
Others	Multiple	Multiple products are shown in the image	266
	Characters	Clothing that is a character collaboration product	309
	Exclusive	Clothing that is written as being for specific users	47

thumbnail images listed in Table 1. In cases where it was difficult to determine whether the character was well-known, we consulted the respondents before labeling.

Table 1 shows that more than half (1,519) of the 3,000 images were captured under floor-standing conditions. We considered the possibility that images captured by sellers on the floor tend to be frequently used as thumbnail images for the tops. However, only a few images labeled with the delivery method and brand logos were used. We believe that these can be included as metadata when setting the product information; therefore, few of them were included as elements in the thumbnail images.

4. Creation of a Classification Model to Identify Product Purchasing Factors

In this study, we identified user purchasing factors by creating four discriminant models: logistic regression, decision tree, gradient boosting, and random forest, all of which aim to predict whether a product is purchased. We created three types of features for the model: metadata only, image labels only, and two components: metadata and image labels. In addition, we revealed the degree to which the accuracy of the same method varies depending on the feature values used in the three patterns, and the feature values that contribute to models with high accuracy. By interpreting these features, we identified the factors that contribute to the purchase of products in online flea markets. In addition, because we use a binary classification model, the objective variables are sold (=1) and displayed (=0).

We used liblinear [18], which is recommended for small amounts of data, as the optimization algorithm for the logistic regression model. In the decision tree, gradient boosting, and random forest models, we selected the maximum tree depth with the smallest difference in correct responses between the training and test data, between 3 and 10, in terms of the amount of training.

In addition, to evaluate each model, we conducted a fivefold cross validation. In the fivefold cross validation, the data are divided into five parts, one of which was the test data, and the rest was the training data. This is repeated five times, and the generalization performance of the model was evaluated by calculating the evaluation values. We used the accuracy, precision, recall, and *F*-score as evaluation indices, which were calculated as the ratio of the predicted result to the actual result being positive or negative. Additionally, we used the area under the curve (AUC), which is based on the receiver operating characteristic (ROC) curve, as an evaluation index and evaluated the performance of each model based on these indices.

4.1. Logistic regression model

The logistic regression model [19] is a type of statistical regression model that predicts the occurrence or nonoccurrence of an objective variable using multiple factors that serve as explanatory variables. Specifically, the likelihood of an event occurring is expressed as a probability using a logit transformation and is classified according to threshold values. Equation (1) indicates the probability of an event occurring as $P(y = 1)$.

$$P(y = 1) = \frac{1}{1 + e^{-\left(w_0 + w_1x_1 + w_2x_2 + \ldots + w_px_p\right)}} \quad (1)$$

The variables are y for the objective variable, x for the explanatory variable, and w for the parameter. We make optimal predictions by setting the parameter values that minimize the loss function where e is the Euler number, which represents the base of the natural logarithm. Table 2 shows the accuracy, recall, *F*-score, and AUC when the logistic regression model showed the highest percentage of accuracy responses after fivefold cross validation. Values in parentheses are the standard deviations of five measurements.

Table 2
Fivefold cross validation score of the logistic regression model

Feature value	Accuracy	Precision	Recall	*F*-score	AUC
Meta	0.652	0.645	0.673	0.659	0.700
	(0.016)	(0.026)	(0.064)	(0.019)	(0.024)
Image Labels	0.653	0.668	0.610	0.638	0.701
	(0.021)	(0.044)	(0.118)	(0.083)	(0.038)
Meta +	0.678	0.680	0.673	0.677	**0.747**
Image Labels	(0.025)	(0.030)	(0.064)	(0.031)	(0.028)

Table 2 shows that the model using both metadata and image labels as features had the highest evaluation values for all the evaluation indices. The AUC values for all feature patterns exceeded 70%, while the values for metadata alone and image labels alone were only approximately 70%, with a slight difference. However, the model using image labels in addition to metadata had an evaluation value of nearly. Thus, we suggest that the use of both metadata and image labels is appropriate for logistic regression models.

4.2. Decision tree model

The decision tree model [20] is a model that learns a hierarchical tree structure composed of conditional expressions with explanatory variables. Table 3 shows the accuracy, recall, *F*-score, and AUC when the decision tree model showed the highest percentage of accuracy responses after fivefold cross validation. Values in parentheses are the standard deviations of five measurements.

Table 3 shows that the highest evaluation values for the accuracy, recall, *F*-score, and AUC are obtained from the model using image labels in addition to metadata. On the other hand, precision is the highest value for the model that uses only image labels as features. In this study, precision indicates the proportion of items predicted by the model as sold that are actually sold. For the decision tree model, it fit that using only image labels as features is suitable for this evaluation index. However, for the other four index, the model that uses both metadata and image labels has higher evaluation values. Therefore, we considered appropriate to use both metadata and image labels as features in the decision tree model.

4.3. Gradient boosting model

The gradient boosting model is an ensemble method that trains multiple weak learners to obtain a more accurate learner, called

Table 3
Fivefold cross validation score of the decision tree model

Feature value	Accuracy	Precision	Recall	*F*-score	AUC
Meta	0.662	0.663	0.657	0.660	0.704
	(0.011)	(0.014)	(0.057)	(0.025)	(0.016)
Image Labels	0.657	0.702	0.550	0.614	0.709
	(0.016)	(0.031)	(0.083)	(0.057)	(0.033)
Meta +	0.675	0.662	0.717	0.688	**0.713**
Image Labels	(0.012)	(0.017)	(0.052)	(0.020)	(0.011)

Table 4
5-fold cross validation score of the gradient boosting model

Feature value	Accuracy	Precision	Recall	*F*-score	AUC
Meta	0.693 (0.029)	0.695 (0.036)	0.690 (0.065)	0.692 (0.033)	0.739 (0.018)
Image Labels	0.667 (0.018)	0.703 (0.027)	0.567 (0.082)	0.628 (0.055)	0.715 (0.033)
Meta + Image Labels	0.718 (0.030)	0.737 (0.039)	0.680 (0.062)	0.707 (0.034)	**0.783** (0.020)

boosting [21]. Boosting obtains a single learner by combining multiple weak learners who are trained by correcting errors from previous weak learners. Table 4 shows the accuracy, recall, *F*-score, and AUC when the gradient boosting model showed the highest percentage of accuracy responses after fivefold cross validation. Values in parentheses are the standard deviations of five measurements.

Table 4 shows that, as with the logistic regression model, the model using image labels in addition to the metadata had the highest evaluation values for all the evaluation values of accuracy, precision, recall, *F*-score, and AUC. Based on the above, we suggest that using both metadata and image labels as features in a gradient boosting model is appropriate. The evaluation index for all features except recall was above 70%, and the AUC was 78%, which is the highest evaluation value among the above two models.

4.4. Random forest model

The random forest is an ensemble method that uses bagging [22]. Bagging uses a bootstrap sample to train a weak learner and performs classification by taking majority votes. Table 5 shows the accuracy, recall, *F*-score, and AUC when the random forest model showed the highest percentage of accuracy responses after fivefold cross validation. Values in parentheses are the standard deviations of five measurements.

Table 5 shows that, similar to the decision tree, the model using both metadata and image labels as features have the highest accuracy, recall, *F*-score, and AUC values. However, the precision is highest for the model using only image labels. Therefore, we considered it appropriate to use both metadata and image labels as features in the random forest model, similar to other models.

4.5. Comparison of model evaluations

We used four methods to create classification models for identifying user purchase factors: logistic regression, decision tree, gradient boosting, and random forest. As a result, for most evaluation indices, models using both metadata and image labels exhibited higher values across multiple evaluation indices. Thus, we considered that adding image labels to metadata allows the creation of more accurate classification models for predicting product purchases in online flea markets. In addition, when comparing the accuracy of using only metadata and image labels as features, the model using only metadata showed higher evaluation values across multiple evaluation indices. This is likely because the metadata includes not only a larger number of features but also features directly related to purchase factors, such as price.

Furthermore, when comparing the four models using both metadata and image labels, the random forest model showed the highest evaluation value for recall and *F*-score, whereas the model using gradient boosting had the highest values for accuracy, precision, and AUC among other evaluation indices. Therefore, the results indicate that the random forest and gradient boost models exhibit relatively high accuracy. In this study, we adopt a gradient boosting model created using metadata and image labels as features to investigate the factors influencing product purchases.

Table 5
5- fold cross validation score of the random forest model

Feature value	Accuracy	Precision	Recall	*F*-score	AUC
Meta	0.672 (0.023)	0.652 (0.033)	0.747 (0.092)	0.693 (0.023)	0.716 (0.022)
Image Labels	0.650 (0.022)	0.672 (0.038)	0.600 (0.121)	0.619 (0.087))	0.678 (0.024)
Meta + Image Labels	0.680 (0.015)	0.660 (0.024)	0.770 (0.112)	0.715 (0.035)	**0.757** (0.024)

5. Results and Discussion on Selected Model Features

Based on the evaluation values of each model, we adopted a gradient boosting model created using metadata and image labels as features to elucidate the factors influencing product purchases. Feature importance, which indicates which features are related to the target variable in the classification, was calculated. Figure 1 shows the feature importance obtained from the gradient boosting model using metadata and image labels as features. Table 6 lists the values of the top 10 features in terms of feature importance obtained by the same model.

From Figure 1 and Table 6, it is evident that features such as the shipping method, price, product condition, and number of likes, which are metadata, are highly important in terms of variable importance. However, the importance of other metadata features, apart from the shipping method, price, and number of likes, is low, while the values for features related to image labels are higher.

Furthermore, we created partial dependence plots for all features to determine whether each feature influenced the classification of the listed or sold items. Partial dependence plots visually depict how specific explanatory variables affect the class classification. In this study, values closer to 1 indicate a contribution to the items sold, whereas values closer to 0 indicate a contribution to the items listed. Figures 2 and 3 show the results of the partial dependence plots for the quantitative variables price and number of likes.

Focusing on the partial dependence plot for price in Figure 2, it is evident that items with extremely low prices tend to be listed rather than sold. As price increases, the probability of an item being sold also increases. In particular, the probability of a commodity being sold exceeds 50% when the price exceeds JPY 1,000, indicating that the commodity tends to sell better. From around JYP 1,500, the probability of being sold increases gradually, peaking beyond JPY 2,000, where the probability is the highest. However, there is a declining trend starting around JYP 2,200, suggesting that prices beyond this point may reduce the likelihood of sales. Based on these observations, when listing tops from fast fashion brands in online marketplaces, setting prices between JYP 1,000 and 2,200, depending on the regular price and item condition, tends to increase the probability of purchase. We consider that items listed within this price range closely match the retail prices of the targeted brand's tops, suggesting a higher likelihood that items listed at these prices are new. Focusing on the partial dependence plot for the number of likes in Figure 3, it is observed that up to three likes, there is an increasing trend toward items being sold,

Figure 1
Feature importance by gradient boosting model

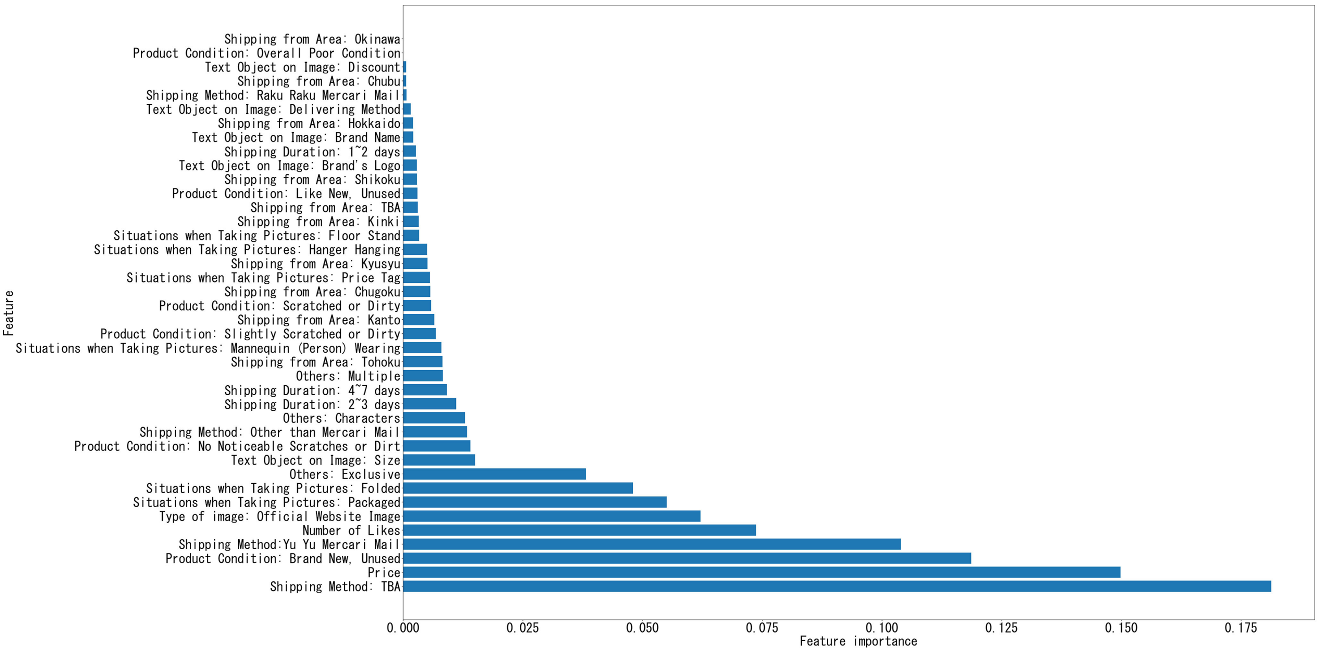


Table 6
Top 10 features by feature importance from the gradient boosting model

Feature value		Feature importance
Meta	Shipping Method: TBA	0.1814
	Price	0.1495
	Product Condition: Brand New, Unused	0.1195
	Shipping Method: Yu Mercari Mail	0.1040
	Number of Likes	0.0736
Image Label	Official Website Image	0.0621
	Packaged	0.0551
	Folded	0.0480
	Exclusive	0.0382
	Size	0.0142

Figure 2
Partial dependence plot of price by the gradient boosting model

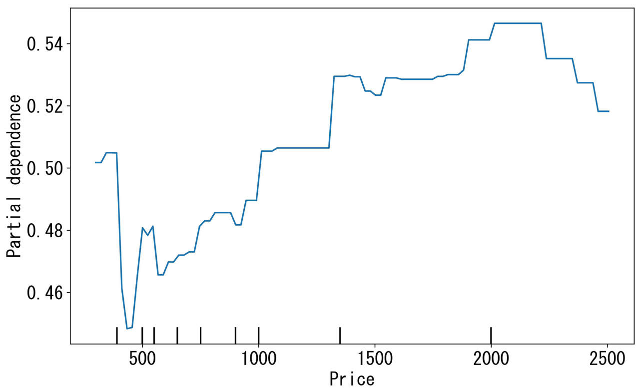
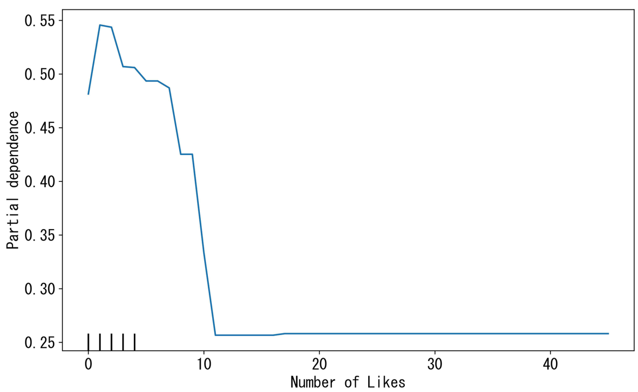


Figure 3
Partial dependence plot of likes by the gradient boosting model



indicating a higher probability of items with some likes being sold. However, for items with four or more likes, the probability of being sold decreases as the number of likes increases, suggesting a trend towards lower sales. Furthermore, at around ten likes, the probability of being sold stagnates at approximately 25%. This indicates that while these items are considered for purchase by users, they may not actually lead to a purchase, potentially because they are listed for a long period and accumulate likes over time. Therefore, products with a certain number of likes are likely to have been listed for an extended period and may have a lower likelihood of being purchased. The results of the partial dependence plots for other features, categorized into five levels of strongly positive, positive, neutral, negative, and strongly negativhyjtkre6tue impacts on the items sold, are summarized in Table 7.

Table 7
The results of partial dependence plots for each feature

Feature value			Impact
Meta	Shipping Method	Yu Mercari Mail	Positive
		Raku Mercari Mail	Neutral
		Other than Mercari Mail	Negative
		TBA	Strongly Negative
	Shipping Origin Region	Hokkaido	Positive
		Tohoku	Positive
		Kanto	Neutral
		Chubu	Neutral
		Kinki	Neutral
		Chugoku	Neutral
		Shikoku	Neutral
		Kyusyu	Positive
		Okinawa	Neutral
		TBA	Neutral
	Number of Days until Shipping	1~2 days	Neutral
		2~3 days	Neutral
		4~7 days	Neutral
	Product Condition	Brand New, Unused	Strongly Positive
		Like New, Unused	Positive
		No Noticeable Scratches or Dirt	Negative
		Slightly Scratched or Dirty	Negative
		Scratched or Dirty	Negative
		Overall Poor Condition	Neutral
Image Label	Text Object on Image	Size	Neutral
		Brand Name	Neutral
		Delivering Method	Negative
		Discount	Negative
		Brand's Logo	Negative
	Situations when Taking Pictures	Floor Stand	Neutral
		Hanger Hanging	Neutral
		Mannequin (Person) Wearing	Strongly Negative
		Folded	Positive
		Packaged	Strongly Positive
		Price Tag	Neutral
	Type of Image	Official Website Image	Strongly positive
	Others	Multiple	Positive
		Characters	Neutral
		Exclusive	Strongly positive

Table 7 shows that the partial dependence plot results for the shipping method, which had the highest variable importance, indicated a significant impact on the items listed rather than those sold. Therefore, products with undefined shipping methods are less likely to be purchased. Because shipping costs vary depending on the method used when shipping is buyer-paid, we suggest that users consider this factor as crucial when selecting products for purchase.

Next, focusing on the results of the partial dependence plot for product condition, it is evident that items labeled as "Brand New, Unused," which showed relatively high feature importance, significantly influence products being sold. On the other hand, product conditions labeled as "Slightly Scratched or Dirty" and "Scratched or Dirty" negatively impact the likelihood of products being sold. This suggests that in a marketplace in which second-hand products are traded, items in near-new conditions are more likely to be purchased.

The results of the partial dependence plot for image labels also show both positive and negative effects on the products sold. First, feature values that significantly influenced items being sold included "Packaged," "Official Website Image," and "Exclusive." For items with a thumbnail image of the packaging, we suggest that the packaging methods of the brands included in this study have an impact. For both brands, many products are packaged in zippered pouches or other packaging materials that serve as price tags. For such products, packaging visually conveys new conditions, leading to a higher likelihood of sales. Additionally, for products using official website images as thumbnails, we consider the listing format of the targeted online flea market to be influential. Users can only view thumbnail images in a square format and their prices in a list format. We assume that official website images that capture a product's features more accurately and appear more esthetically pleasing might appeal more to users at a glance. We suggest that the labeling of items as "Exclusive" has an effect on the number of items sold since the buyer is already determined at the time of the sale.

We also consider that the labeling of wearing a mannequin (person), which has a significant negative effect, has a negative effect because the amount of data with this label is small, and the products are not new or unused when worn by a person.

6. Conclusion

In this study, we revealed the purchasing factors in an online flea market by conducting a binary classification of sold and unsold items.

Specifically, we used features such as metadata (price, shipping method, etc.) and manually labeled images from product thumbnails. Previous research [7, 8] used image features using convolutional neural network models in which image features are considered in bulk. On the other hand, in this study, we analyzed image features by dividing them into multiple components from the perspective of online flea markets. We consider that this leads to a more detailed understanding of image features in online flea market purchases. Specifically, models were created using three patterns: metadata only, image labels only, and metadata and image labels alone. We created 12 models using four methods: logistic regression, decision tree, gradient boosting, and random forest. Consequently, the gradient boosting model using both metadata and image labels as features achieved high evaluation values.

Furthermore, we revealed the potential factors in the purchase of products in online flea markets by creating feature importance and partially dependent plots using a gradient boosting model with both metadata and image labels, which obtained high evaluation values. Factors significantly affecting product sales included prices in the range of JYP 1,000–2,200 and products in new and unused conditions. Regarding image labels, we found that products using the official homepage or an image of the product being packaged as a thumbnail tended to be sold more often. Few previous studies have focused on thumbnails to elucidate consumer decision making, and through this study, we revealed that the characteristics of thumbnail images contribute to consumers' purchasing behavior.

In the future, we suggest that more detailed results can be obtained by creating a model that considers not only thumbnail images but also other product images and by expanding the data. In addition, since the target of this study was fast fashion brands, we believe that official website images and packaging affected the image labels sold. Therefore, we believe that different features affect the sales results using other categories. Furthermore, manual labeling was used to extract image features in this study, but we will attempt to extract information from images and text using neural networks in the future, which have been widely used in recent years [23].

Funding Support

In this paper, we used "Mercari Dataset" provided by Mercari, Inc. via IDR Dataset Service of National Institute of Informatics. We express my heartfelt gratitude to Mr. Hayato Iwasaki for his invaluable assistance in advancing this research. Furthermore, this work was supported by JSPS KAKENHI Grant Number 21K13385.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Author Contribution Statement

Emi Iwanade: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Yoshihisa Shinozawa:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Writing–original draft, Writing – review & editing, Visualization. **Kohei Otake:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

References

[1] Daiwa Institute of Research Ltd. (2021). *2021 market research on electronic commerce*. Retrieved from: https://www.meti.go.jp/press/2022/08/20220812005/20220812005-i.pdf

[2] Hellianto, G. R., Suzianti, A., & Komarudin (2019). User experience modeling on consumer-to-consumer (c2c) e-commerce website. In *IOP Conference Series: Materials Science and Engineering*, *505*. https://doi.org/10.1088/1757-899X/505/1/012080

[3] Avcı, İ., & Yıldız, S. (2023). The effects of online second-hand product purchase motivations on sustainable consumption behavior and online repurchase intention in consumer-to-consumer electronic commerce (C2C) and the mediating role of electronic word-of-mouth marketing. *Pazarlama Ve Pazarlama Araştırmaları Dergisi*, *16*(1), 137–164.

[4] Moriuchi, E., & Takahashi, I. (2023). An empirical study on repeat consumer's shopping satisfaction on C2C e-commerce in Japan: The role of value, trust, and engagement. *Asia Pacific Journal of Marketing and Logistics*, *35*(3), 560–581. https://doi.org/10.1108/APJML-08-2021-0631

[5] Okamoto, T., Yatsuhashi, J., & Mizutani, N. (2020). Consumers' priorities for online Flea market. In *Proceedings of the 7th Multidisciplinary in International Social Networks Conference and the 3rd International Conference on Economics, Management and Technology, 6*, 1–6. https://doi.org/10.1145/3429395.3429400

[6] Sihvonen, J., & Turunen, L. L. M. (2016). As good as new-valuing fashion brands in the online second-hand markets. *Journal of Product & Brand Management*, *25*(3), 285–295. https://doi.org/10.1108/JPBM-06-2015-0894

[7] Wang, Y., Guo, Y., & Song, J. (2018). Using image-based and text-based information for sales prediction: A deep neural network model. In *Americas Conference on Information Systems 2018: Digital Disruption, AMCIS 2018* (Americas Conference on Information Systems 2018: Digital Disruption, AMCIS 2018). Association for Information Systems.

[8] Zhong, H., & Wang, Y. (2021). A product recommendation method by analyzing images and reviews in e-commerce. In *2021 IEEE 10th Global Conference on Consumer Electronics*, 934–935. https://doi.org/10.1109/GCCE53005.2021.9621927

[9] Valavan, M., & Rita, S. (2023). Predictive-analysis-based machine learning model for fraud detection with boosting classifiers. *Computer Systems Science and Engineering*, *45*(1), 231–245. https://doi.org/10.32604/csse.2023.026508

[10] Wu, H., & Li, B. (2022). Customer purchase prediction based on improved gradient boosting decision tree algorithm. In

Proceedings of the 2022 2nd International Conference on Consumer Electronics and Computer Engineering, 795–798. https://doi.org/10.1109/ICCECE54139.2022.9712779

[11] Ghosh, S., & Banerjee, C. (2020). A predictive analysis model of customer purchase behavior using modified random forest algorithm in cloud environment. In *Proceedings of the 2020 IEEE 1st International Conference for Convergence in Engineering*, 239–244. https://doi.org/10.1109/ICCE50343.2020.9290700

[12] Zhong, Z., Yuan, X., Liu, S., Yang, Y., & Liu, F. (2021). Machine learning prediction models for prognosis of critically ill patients after open-heart surgery. *Scientific Reports*, *11*(1), 3384. https://doi.org/10.1038/s41598-021-83020-7

[13] Sato, R., & Tamura, R. (2019). Study on thumbnail images and titles selected by viewers in YouTuber's video. *International Journal of Affective Engineering*, *18*(1), 139–145. https://doi.org/10.5057/jjske.TJSKE-D-18-00038

[14] Zhang, S., Aktas, T., & Luo, J. (2021). Mi YouTube es Su YouTube? Analyzing the cultures using YouTube thumbnails of popular videos. In *2021 IEEE International Conference on Big Data*, 4999–5006. https://doi.org/10.1109/BigData52589.2021.9672037

[15] Miyamoto, Y. (2022). An approach to classify thumbnail images on video sites by the number of accesses. In *37th International Technical Conference on Circuits/Systems*, 194–197. https://doi.org/10.1109/ITC-CSCC55581.2022.9894950

[16] Kozu, H., Saito, K., Nakamura, S., Kondo, R., & Tsujimura, Y. (2024), A study on sensibility evaluation of thumbnails on online video sharing services for tourism targeting Japanese elderly people. *Journal of Global Tourism Research*, *9*(1), 13–21. https://doi.org/10.37020/jgtr.9.1_13

[17] Mercari, Inc. (2023). *Mercari Dataset*. Informatics Research Data Repository, National Institute of Informatics. https://doi.org/10.32130/idr.17.1

[18] LIBLINEAR – A Library for Large Linear Classification. Retrieved from: https://www.csie.ntu.edu.tw/~cjlin/liblinear/

[19] Cox, D. R. (1958) The regression analysis of binary sequences. *Journal of the Royal Statistical Society*, *20*(2), 215–232. https://doi.org/10.1111/j.2517-6161.1958.tb00292.x

[20] Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*, 81–106. https://doi.org/10.1007/BF00116251

[21] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5), 1189–1232. http://www.jstor.org/stable/2699986

[22] Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1023/A:1010933404324

[23] Lu, S., Ding, Y., Liu, M., Yin, Z., Yin, L., & Zheng, W. (2023). Multiscale feature extraction and fusion of image and text in VQA. *International Journal of Computational Intelligence Systems*, *16*(1), 54. https://doi.org/10.1007/s44196-023-00233-6

How to Cite: Iwanade, E., Shinozawa, Y., & Otake, K. (2025). Identifying Purchasing Factors in Online Flea Markets Considering Thumbnail Images. *Journal of Data Science and Intelligent Systems*. https://doi.org/10.47852/bonviewJDSIS52024073