ISASE 2020

Recommender system based on personal kansei evaluation tendency

Yuya KONDO*, Hiroshi TAKENOUCHI ** and Masataka TOKUMARU ***

* Graduate School of Kansai University, 3-3-35 Yamate-cho, Suita-shi, Osaka 564-8680, Japan

** Fukuoka Institute of Technology, 3-30-1 Wajiro-higashi, Higashi-ku, Fukuoka 811-0295, Japan

*** Kansai University, 3-3-35 Yamate-cho, Suita-shi, Osaka 564-8680, Japan

Abstract: In the present paper, we propose a recommender system that considers the current user behavior tendencies to evaluate relevant contents. In, the proposed model, the user preference factors are estimated based on the collected kansei evaluations and image features of recommended contents. The features are acquired from the image of the content by using a deep neural network, and the user evaluation for each feature is estimated by applying the gradient boosting decision tree. In addition, we analyze the relationship between the user evaluation and the image feature, and search for the users with the similar evaluation tendency with regard to changes in the feature. Based on the obtained results, we recommend the individuals that are presumed to be highly rated by other users. In addition, we estimate the features that are the factors of the user preference corresponding to recommendation results and present the common features as images. The effectiveness of the proposed model is verified by performing a simulation using the user evaluation data.

Keywords: Recommender System, Neural Network, Gradient Boosting, Decision Tree

1. INTRODUCTION

In recent years, the number of contents in the EC site and entertainment online streaming services has been enormous. Therefore, it is difficult for users to identfy the best contents among all available candidates by themselves. Many services employ recommendation systems to suggest relevant contents based on the user evaluation results and enable searching for contents that may be interesting for users[1]. Collaborative filtering(CF)[2] and content-based filtering are the search methods commonly used in recommender systems. However, these filtering methods difine recommendations based on the evaluation of users and the degree of matching the content attributes. Therefore, they do not incorporate the information corresponding to the kansei evaluation, namely, which part of the content the users mostly pay attention to and what impression they have. In addition, recommendation of contents in the currently available systems does not provide users with explanation on why particular contents are recommended. These problems may affect the confidence of users while employing recommender systems, as well as satisfaction with the recommendation accuracy.

Therefore, we propose a recommendation system that is based on the tendencies of user kansei evaluation. In the system, we examine the estimation of the user preference factors based on the collected kansei evaluation and the image features of the recommended contents. The system is applied to estimate the user evaluation based on the features extracted from the content images. In addition, we analyze the relationship between the user evaluation tendencies and features, and search for users with the similar kansei evaluation tendency. Furthermore, the common features of the user's favorite content are identified by restoring the image from the feature based on the obtained sensitivity tendency. The effectiveness of the proposed system is confirmed by performing a simulation.

2. RECOMMENDER SYSTEM MODEL

The proposed model consists of three parts: a deep neural network (DNN) used to extract features from content images, decision tree that learns the user kansei evaluation, and a filtering part of the kansei evaluation tendency.

2.1 Feature extractor and image restoration unit

According to previous research [3], it is possible to employ an autoencoder, which is a kind of DNN, aiming to extract features from images for emotional information retrieval. In our proposed model, we use a convolutional autoencoder, which provides better results in extracting



Figure 1: Configuration of the autoencoder

image features.

Fig. 1 represents the structure of the convolutional autoencoder implemented in the proposed model. The encoder shown in Fig. 1 (a) is used to compress a 128 * 128 pixel color image into eight-dimensional features. In addition, the input image can be reconstructed from the eight-dimensional features by using the decoder. Based on the user evaluation and the features extracted by the encoder, the user emotional evaluation tendency is estimated. The extracted features are represented by eight real numbers for each image. In addition, an input image can be reconstructed based on the eight-dimensional features by using the decoder, as shown in Fig. 1 (b). In the present study, we attempt to visualize the features of kansei evaluation by restoring images only from specific features.

2.2 Gradient boosting

Gradient boosting is boosting algorithm proposed by Friedman [4]. Gradient boosting typically applies the fixed-size decision trees (especially classification and regression trees) as basic learners. For this specific case, Friedman proposes a modification of the gradient boosting method that improves the quality of fitness of each elementary learner. The generated decision tree is called Gradient boosting decision tree (GBDT).

The goal of this algorithm is to find an approximation $\hat{F}(x)$ of the function F(x) such that to minimize the expected value of a particular loss function L(y, F(x)). General gradient boosting at the *m*-th step fits the decision tree $h_m(x)$ to the pseudo residuals. Let J_m be the number of leaves. The tree divides the input space into J_m unconnected regions $R_{1m}, \dots R_{J_mm}$ and predicts a constant value for each region. Using indicator notation, the output of $h_m(x)$ corresponding to input x can be written as a sum as in Eq. (1), where b_{jm} is the value predicted in

region R_{jm}.

$$h_m(x) = \sum_{j=1}^{J_m} b_{jm} \mathbf{1}_{R_{jm}}(x)$$
(1)

Then, the coefficients b_{jm} are multiplied by values γ_m selected by applying the line search to minimize the loss function L(y, F(x)), and the model is updated as per Eq. (2):

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x),$$

$$\gamma_m = \arg \min_{\gamma} \sum_{i=1}^n L(y_i, F_{m-1}(x_i) + \gamma h_m(x_i)) \quad (2)$$

Friedman proposes a modification of this algorithm to select individual optimal values γ_{jm} for each region of the tree, rather than a single γ_m for the entire tree. It is referred to as the modified algorithm "TreeBoost". The coefficients b_{jm} obtained as a result of the tree fitting procedure are replaced, and the model update rule is as shown in equation (3).

$$F_{m}(x) = F_{m-1}(x) + \sum_{j=1}^{J_{m}} \gamma_{jm} \mathbf{1}_{R_{jm}}(x),$$

$$\gamma_{m} = \arg \min_{\gamma} \sum_{x_{i} \in R_{jm}}^{n} L(y_{i}, F_{m-1}(x_{i}) + \gamma) \quad (3)$$

In the present study, we estimate user kansei evaluation by applying multiclass classification based on GBDT.

2.3 Collaborative filtering by evaluation tendency

To recommend the content that may be relevant for users, the recommender system searches for users with the similar kansei. In the present study, user ratings are collected in ten levels. The similarity of the kansei tendency means that the change in evaluation is large relatively to the change in the characteristics used for evaluation judgment. Therefore, by comparing the consistency of the evaluation differences between users with regard to each image, the users with the similar evaluation tendency are can be found.

The system calculates the difference between the evaluation value corresponding to the image evaluated by the user and those of other users. If there is a high degree of similarity in evaluation trends among users, the difference in evaluation values corresponding to these users is almost constant. However, it is unclear whether the user kansei evaluation is consistent for all images. Therefore, in the present study, the evaluation up to the absolute value of 1 is considered as indicating no significant difference in impression of the users. According to the above points, the system calculates difference the between user ratings for each image and the average of user ratings for all images. To examine the evaluation error corresponding to two different users, the ratio of images in the range from -2 to 2 is defined as the similarity of the evaluation tendency. The estimated evaluation value $\dot{E}(x)$ of the user u_1 for image x_i is expressed by Eq. (4) using the average point $Avg(u_1, u_2)$ of the difference between the user u_2 and the similarity $S(u_1, u_2)$ of the tendency.

$$\dot{E}(u_1, x_i) = E(u_2, x_i) + Avg(u_1, u_2) \times S(u_1, u_2)(4)$$

Based on the user evaluation in which the system has determined that the similarity of the evaluation tendency is high, the corresponding image for which the high evaluation has been estimated is recommended to the user.

3. VERIFICATION USING THE USER DATA

3.1 Collect user evaluations

In the present study, we extract features from images and estimate kansei evaluation. Therefore, it is preferable that the user can easily define the preference of the content based on the image. To confirm the effectiveness of the proposed model, we use T-shirt images as evaluation targets. Fig. 2 represent the examples of T-shirt images used to perform the evaluation. With the range having 10 as "like" and 1 as "dislike", ten levels of preference evaluations for 1131 T-shirt images were collected from 18 men and women in their 20s.

3.2 Autoencoder and GBDT parameters

Tab. 1(a) presents the learning parameters of the encoder part of the autoencoder used to extract eightdimensional features from 128 * 128 pixel images. Tab. 1(b) provides the parameters of GBDT for estimating the



Figure 2: Example of the experimental T-shirt images

Table 1: Condition of the experiment(a) Autoencoder learning parameters

Number of classification classes	10
max round	100
early stopping rounds	10
Test data ratio	0.8
Learning rate	0.1
Evaluation index	Multiclass logloss
max_depth	6

(b) GBDT learning parameters

epoch	2000
batch size	32
Activation function in	ReLU
encoder	softplus
Activation function in	ReLU
decoder	sigmoid
loss function	MSE

user preference evaluation based on the extracted eightdimensional features.

In Tab.1 (a), "epoch" donates the number of training iterations. The batch size is the size of the entire dataset used as is input to the autoencoder. All data are learned by being dividing into the subsets of the batch size. The activation function of the encoder uses the rectified linear unit (ReLU) function for convolution layers and the softplus function in the last layer of the dence layer. The activation function of the decoder employs the ReLU function in the convolutional layer and the sigmoid function in the last layer of the reconstruction. The loss function is the mean square error. The autoencoder learns to minimize the error between the input image and the restored one. In Tab.1 (b), the number of classification classes is ten similarly as at the user evaluation stage. Max round donates the maximum number of boosts, and early stopping rounds mean the index to stop learning if the error does not improve for a certain number of consecutive rounds. In this experiment, there are ten rounds. The evaluation index is multiclass logloss and learning progresses to generate a tree with the minimum index corresponding to the test data.

3.3 Results and discussions

Fig. 3 shows the distribution of evaluation points for all users. A rating of 7 or more is 13% of the total. Therefore, users did not like this dataset image much. The effectiveness of the proposed system has been verified for

one of the selected users. Fig. 4 represents the decision tree defined by using GBDT. Based the constructed decision tree, the system extracts the rules that are leaf nodes of the highly evaluated images. Then, the rule of evaluation 7 is extracted with the feature amount $x^2 < x^2$ 120.937096 and $x4 \ge 71.4925623$. In addition, the evaluation is predicted based on the evaluation of the user with the closest evaluation tendency among the other 17 users. Among all the images which evaluated this time, about 20% of the images scored 7 or more. Therefore, Fig. 5 represents the error between the predicted values corresponding to images with 7 or more points and the users evaluations for both decision tree and evaluation tendency filter. Finally, based on the extracted rules, Figure 6 represents the images reconstructed from image features x2 and x4 which are classified into leaf nodes of the extracted rules.

In GBDT, the error corresponding to the high evaluation image is large, and the leaf node with the maximum





Figure 3: Distribution of evaluation points for all users

Figure 4: GBDT constructed based on the selected user evaluation and T-shirt features



Figure 5: Average of evaluation error for the selected user of GBDT and kansei tendency filter



Figure 6: Images restored from features x2 and x4

evaluation value is 7. This is because most of the T-shirt images used for evaluation have not been liked by the subject, and only 20% of the evaluation images with 7 or more points are evaluated. In addition, the commonality based on the restored image has not been confirmed. The diversity of the T-shirt images used in this experiment has been too large with respect to the number of images, so that the features may not have been successfully restored.

However, as the average error due to filtering for a particular user is smaller than two in the case when the number of the high-evaluation candidates is small as in this experiment, we can conclude that prediction based on the similarity of the evaluation tendency is more effective than the rule defined according to the decision tree.

4. CONCLUSION

In the present research, we proposed the recommender system composed of an autoencoder, GBDT, and kanseibased filtering. As a result, the method implying extraction of highly evaluated rules and estimation of preference factors according to the restored image based on the rules was found inappropriate. However, as the average error due to filtering was less than two, making recommendation based on the sensitivity evaluation tendency may be effective. In our future work, verification using a larger data set including the large amount of user data and conduct a study on verbalization of restored image features.

REFFERENCES

- G. Linden, B. Smith and J. York, "Amazon.com recommendations: item-to-item collaborative filtering," in IEEE Internet Computing, Vol. 7, No. 1, pp. 76-80, 2003.
- [2] M. Balabanovic Y. Shoham. Fab: Content-based, collaborative recommendation.Communications of the ACM, Vol. 40, No. 3, pp. 66-72, 1997.
- [3] S. Ota, H. Takenouchi, M. Tokumaru, "Kansei clothing retrieval system using features extracted by autoencoder", 2017 IEEE Symposium Series on Computational Intelligence (SSCI) Proceedings, pp. 281-287, 2017.
- [4] J. H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine", The Annals of Statistics, Vol. 29, No. 5, pp. 1189-1232, 2001.