International Journal of Electrical and Computer Engineering (IJECE)
Vol. 14, No. 2, April 2024, pp. 1711~1719
ISSN: 2088-8708, DOI: 10.11591/ijece.v14i2.pp1711-1719

Developing a restaurant recommended system via the Vietnamese food image classification

Viet Hoang Pham[1], Anh Thai Nguyen[1,2], Bao The Phung[1], Truong Ho-Viet Phan[1]
[1]Faculty of Information Technology, Van Lang University, Ho Chi Minh City, Vietnam
[2]School of Engineering and Technology, Asian Institute of Technology, Pathum Thani, Thailand

Article Info

Article history:

Received Aug 16, 2023
Revised Oct 6, 2023
Accepted Nov 4, 2023

Keywords:

Alternative least square
Apache Spark
Image classification
Recommendation system
ResNet

ABSTRACT

A recommendation system is a system that recommends products and services to users based on daily online searching habits. The recommender system is applied in many fields such as job searching, health care, education, music, and tourism. However, few studies have combined computer vision and collaborative filtering to build a restaurant recommendation system in the tourism sector. In this study, we presented a solution to build a restaurant recommendation system through Vietnamese food image classification. First, we used ResNet-34 which is a variant of the convolutional neural network to classify Vietnamese food images. Then, the system applied the alternative least square technique in matrix factorization and Apache Spark in distributed computing to train the restaurant location dataset. The output was the most relevant restaurant places list to show many choices to users. The experimental datasets included the Vietnamese image and the restaurant location datasets that were collected from kaggle.com and foody.vn websites. For image classification task evaluation, we compared ResNet-34 to variants of ResNet. For the restaurant recommendation task evaluation, we compared alternative least squares with k-nearest neighbor. The comparison results show that the proposed solution is better than traditional popular models.

Corresponding Author:

Truong Ho-Viet Phan
Faculty of Information Technology, Van Lang University
69/68 Dang Thuy Tram, 13 Award, Binh Thanh district, Ho Chi Minh City, Vietnam
Email: truong.phv@vlu.edu.vn

1. INTRODUCTION

A recommendation system is a system that suggests products, services, and utilities that suit the needs of users. Currently, the applications of recommendation systems to specific fields have become more and more popular [1]–[4]. However, few systems recommend culinary places in tourism. Current research in the world in general and in Vietnam in particular typically applies the recommendation system on e-commerce [5], [6]. Therefore, the combination of computer vision and the recommendation system in tourism have opened a new research direction for the recommendation system as well as improving the life quality for people. Besides, the use of distributed computing instead of single computation like other traditional recommendation systems also brings a big breakthrough in reducing data processing time, improving accuracy, and improving data quality [7], [8].

The study of recommendation systems has a rather long history. From using different methods and algorithms to come up with applications that help users more easily reach their needs [9]–[12]. Hussien *et al.* [13] solved the challenges of recommendation systems such as cold start, lack of user matrix, scalability. They proposed a system that used statistical and analytical methods to calculate some features (customer behavior) to build a recommendation system that provides recommendations close to customer preferences. Mauro *et al.* [14] integrate service providers on a recommendation system to introduce different stages of item outcome. They combined the item's attributes and the consumer's perception of the item output. In particular, they took service journey maps for the service-based item and user profiles and proposed a two-level visual model for users' selection. The higher level presented the concentration of consumers about items and supported the prediction of potential items. The lower-level exploited property data of items. The model of Mauro has not handled a great number of users and only used a dependency parser to analyze the aspects of review text, so it had to match aspects with ad hoc dictionaries. Kanout [15] proposes the CoCl model, a context clustering-based recommender system, to find relevant recommendations. He utilized contextual information and K-mean algorithm to create a user-item matrix. He carried out three stages. In the first stage, the CoCl model clustered the ratings and user-item matrix based on contextual information. In the second stage, the CoCl model assembled the user-item matrix from the first stage. In the final stage, the CoCl model leveraged collaborative filtering to recommend the missing preference for users. Kanout's model is not interested in a real-time context-aware recommender system. Also, he experimented on a small dataset LDOS-CoMoDa that consisted of 200 users, 2,296 ratings, and 4,138 movies. Tripathi *et al.* [16] built a restaurant recommender system to assist users in predicting the ratings between users and restaurants. They implemented the k-nearest neighbors' algorithm and multiclass support vector machine to compute the similarity between users and others. Tripathi used a data source from Yelp with a user rating is 5,245,126 records, a business count is 174,576, and a user account is 1,326,100. Tripathi's model is not interested in environment rating, services, and sentiment analysis.

However, few studies have combined computer vision and collaborative filtering to produce a recommender system in a certain domain such as tourism. So, this paper proposes a solution to build a restaurant recommendation system through Vietnamese dish image classification. Specifically, this study used a variant convolutional neural network to recognize Vietnamese food images as input. Then, the model took the image input and ratings as a restaurant-item matrix. Finally, the model applied alternating least squares (ALS) to predict the top 10 restaurant locations for tourists. The system was operated on the web to help users have more experience. The main contributions are summarized as follows: i) using the ResNet model to recognize Vietnamese food from user input, ii) proposing a food-restaurant matrix and applying the alternating least squares to predict the top 10 suitable restaurants for users.

2. METHOD

In this section, we described our system in detail as shown in Figure 1. Our system has three components: image and restaurant dataset, image classification and recommendation, and relevant restaurant list. In the image and restaurant dataset module, we presented the way that we collected information and images from kaggle.com and foody.vn websites. In the image classification module, we presented the ResNet model, which was the improved variant of the convolutional neural network, to classify an image into food categories. In the final module, we leveraged the ALS algorithm of the recommendation system to list relevant restaurants with input images.

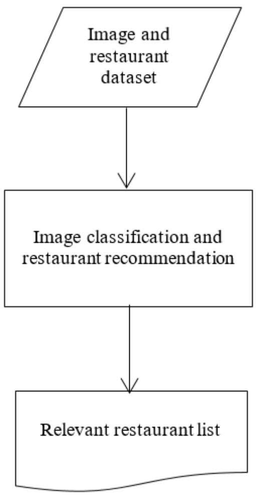

Figure 1. Overview diagram of restaurant recommendation system

2.1. Image and restaurant dataset

2.1.1. Image dataset

For research purposes, we chose two free Vietnamese food datasets on kaggle.com: Vietnamese_food and Vietnamese foods dataset as shown in Table 1 [17], [18]. The first dataset consisted of 10 categories, 5,510 image files; the second dataset consisted of 10 categories, 2,865 image files. Because of some duplicate categories, we merged both of their files into one category. Besides, the images of each dataset had a different size, so we manually resized all images into 244×244 dimensions to fit the input parameters of the training model. This dataset was divided into two parts: 80% for the training set and 20% test set.

Table 1. Summarize the number of images for 18 categories crawled from kaggle.com

#	Categories	Vietnamese_food	Vietnamese foods dataset	Total
1	Banh beo	451	0	451
2	Banh chung	354	0	354
3	Banh mi	935	300	1,235
4	Banh pia	310	0	310
5	Banh xeo	821	300	1,121
6	Ca kho to	473	0	473
7	Canh chua	577	0	577
8	Com tam	659	0	659
9	Pho noodle	564	0	564
10	Xoi xeo	366	0	366
11	Banh bo	0	300	300
12	Banh bot loc	0	300	300
13	Banh canh	0	232	232
14	Banh da lon	0	300	300
15	Banh khot	0	300	300
16	Banh tai heo	0	300	300
17	Banh tieu	0	286	286
18	Banh trung thu	0	247	247
		5,510	**2,865**	**8,375**

2.1.2. Restaurant dataset

This dataset supported our model to establish a user-restaurant matrix. The foody.vn is a food delivery application that is very popular in Vietnam. We crawled the information of restaurant locations and ratings on foody.vn website [19]. On this website, we can choose any Vietnamese food dish by typing the food name. Table 2 shows the summary of restaurant locations for each food category.

Table 2 shows the number of restaurant locations for 18 Vietnamese dish categories with a total was 14,413 addresses. All restaurant information was saved in a text file. The structure of this file included four columns such as food name, restaurant name, restaurant address, and rating as shown in Table 3. In Table 3, the food name column was set the same as the food category names. The restaurant name and addresses were used to recommend suitable restaurants according to the food name that was input from users. The fourth column is used to evaluate the quality of a restaurant.

Table 2. Summarize the number of restaurant locations for 18 food categories crawled from foody.vn

#	Categories	Restaurant locations
1	Banh beo	656
2	Banh chung	506
3	Banh mi	1,336
4	Banh pia	444
5	Banh xeo	1,173
6	Ca kho to	676
7	Canh chua	825
8	Com tam	942
9	Pho noodle	807
10	Xoi xeo	523
11	Banh bo	744
12	Banh bot loc	720
13	Banh canh	964
14	Banh da lon	1,140
15	Banh khot	835
16	Banh tai heo	795
17	Banh tieu	686
18	Banh trung thu	641
		14,413

Table 3. The structure of the text file saved restaurant locations

Food name	Restaurant name	Restaurant address	Rating
Banh beo	Quan Trang – Mi Quang & Banh Beo	68/6 Lu Gia, P15, Q11, HCM	6.4
Banh bot loc	Banh Bot Loc Co Be	327/28/9 Quang Trung, Go Vap	7.5
Banh can	Dalat foods – banh trang nuong	58/8 Dong Nai, P15 Q10	7.0
Banh canh	Banh canh cua	87 Tran Khac Chan, Q1	6.8

2.2. Image classification and recommendation

In this section, we described the process of classifying food images. We also recommend suitable restaurants as shown in Figure 2. In this process, there were five stages: i) the input was a Vietnamese food image, ii) the ResNet-34 model took the image and got the image feature through blocks of 33 layers, iii) the final layer of the ResNet-34 was a full connection and softmax function to predict the food category name, iv) the food category name become the input of food-restaurant matrix, and v) the ALS algorithm was applied to output the relevant restaurant list.

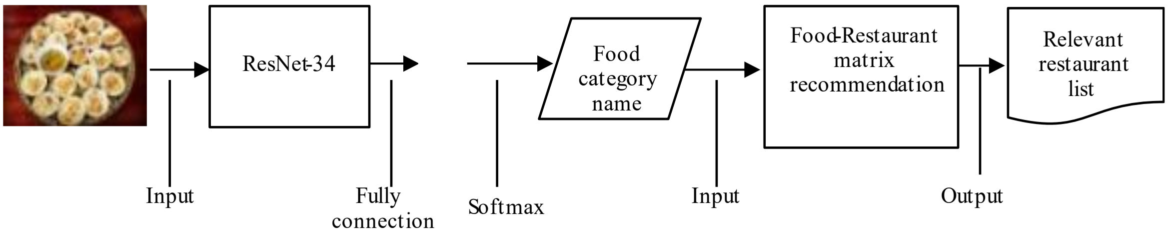

Figure 2. The process of image classification and recommending relevant restaurants

2.2.1. Food image classification

Convolutional neural network (CNN) is one of the deep learning models for image recognition [20], [21]. The CNN model has many variants which typically are residual network (ResNet) models. ResNet is a convolutional neural network designed to work with hundreds or thousands of convolutional layers [22]. Recently, ResNet has had many variants with different layers such as ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152. ResNet solved the vanishing gradient problem which caused the increasing error rate when adding more layers. ResNet introduced the residual block with a skip connection technique. This technique linked the activations of a layer with the next layer by skipping intermediate layers between them to form residual blocks.

In this study, we chose the ResNet-34 to train the food image classification task. Specifically, we randomly selected 80% of food images as input data for the ResNet-34 model. First, the ResNet-34 executed a convolutional operation by using a kernel size of 7, and a feature map size of 64. The output size of the convolutional operation was 112×112. Secondly, the ResNet-34 used batch normalization which was an element-wise operation. Thirdly, the model applied the max pooling operation with a stride of 2. The three above operations formed a block. A layer of the ResNet-34 had many blocks. The final layer was a full connection operation. A softmax function was applied to the output of the ResNet-34 to predict food names.

2.2.2. Restaurant recommendation

The current development directions of recommendation systems are collaborative filtering, content filtering, and hybrid recommendations [11]. In this module, we chose the collaborative filter because of two reasons. First, this technique gives suitable suggestions based on the user's interests without knowing about that product. Second, this method is suitable for large datasets. We denote F_m as the food set with m as the number of food categories, L_n as the restaurant location set with n as the number of restaurant locations as shown in Figure 3. A matrix $A_{m\times n}$ represented the food-restaurant rating matrix.

Next, we used matrix factorization to decompose the matrix $A_{m\times n}$ into small matrices. Afterward, we applied alternate least squares (ALS) to optimize small matrices [23]–[26]. ALS is an iterative optimization process using least squares for every loop step. At every step, the ALS model achieves minimum error rates when compared to our original data. In this study, we split the $A_{m\times n}$ into foods – hidden features matrix C and hidden features – restaurants D.

$$C = F\times h \tag{1}$$

$$D = h \times L \tag{2}$$

$$A \approx C \times D \tag{3}$$

where, F is food matrix, h is hidden feature, and L is restaurant matrix

Initially, we assigned random values in C and D. Next, we used least squares with many loops until the product of C and D obtained the best values that were equal to matrix A. We used the ALS method as introduced above with Apache Spark to train the model efficiently and quickly because Spark was able to handle big data. We split the data into 2 parts: 75% for training and 25% for testing. When the model was trained, our data table had a new column named "prediction". This was the column that returned the suggested results to users. The higher the "prediction" was, the more relevant the suggestion was to the users. Because the training process was as time-consuming as training an image classification model, we also stored this model on a local machine and reused it when we requested.

	L_1	L_2	…	L_{n-1}	L_n
F_1	r_{11}	r_{12}	…	r_{1n-1}	r_{1n}
F_2	r_{21}	r_{22}	…	r_{2n-1}	r_{2n}
…	…	…	…	…	…
F_{m-1}	r_{m-11}	r_{m-12}	…	r_{m-1n-1}	r_{m-1n}
F_m	r_{m1}	r_{m2}	…	r_{mn-1}	r_{mn}

Figure 3. The food-restaurant matrix represented the relationship between the food and restaurant dataset

3. RESULTS AND DISCUSSION

In this section, we experimented with our Vietnamese food and restaurant address dataset. To evaluate the performance of our model, we compared it with the series of variant ResNet models for image classification. For big data processing, we compared our model using Apache Spark with MS SQL Server.

3.1. Setting

The restaurant recommendation system was designed and built on a personal computer. Its configuration consisted of Intel's CPU i7-8565U series consisting of 4 cores and 8 threads with the processor's-based frequency of 1.80 GHz, 16 GB RAM. We used fastai framework to implement the ResNet model with epochs = 20, layers = {18, 34, 50, 101}.

3.2. Dataset

This study was conducted based on collecting a data set of 8,375 images and 18 popular dishes in Vietnam. We divided these images into two folders train and test. Each folder had 18 food categories. In addition, there was a dataset of 5 million records about restaurants in Vietnam collected on foody.vn. We only crawled 14,413 restaurants but for big data handling purposes, we duplicated these restaurants up to five million.

3.3. Metrics

To evaluate how good our recommendation system was, we used mean absolute error (MAE), and root mean square error (RMSE). Mean absolute error represents the absolute value for each error in rating prediction. The MAE score is lower and the output model is better. Root mean square error has a stronger penalty for prediction when output values and true values have a far distance and a weaker penalty when output values and true values have a close distance. The RMSE score is lower and the output model is better.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \tag{4}$$

where, y_i is the predicted value, x_i: is the actual value, n is total number of prediction values

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{n}} \tag{5}$$

where, x_i is the actual value, $\hat{x}_i$ is the predicted value, and N is the total number of predicted values

3.4. Comparison of Vietnamese food image classification with variants of ResNet

We conducted the food image classification on variant ResNet models such as ResNet-18, ResNet-34, ResNet-50, and ResNet-101. ResNet solved the vanishing gradient problem which caused the increasing error rate when adding more layers. The second reason why we chose this model was that it won the large-scale visual recognition challenge (ILSVRC) competition in 2015 with error rates of only 3.5%. Table 4 shows the result of each ResNet model after training the Vietnamese food dataset.

Table 4. The performance comparison of variant ResNet models

Models	Time (Hours)	Train loss	Valid loss
ResNet-18	3	0.007	1.338
ResNet-34	4	0.006	1.218
ResNet-50	5	0.005	1.164
ResNet-101	11	0.003	1.115

In Table 4, we saw that ResNet architectures with different numbers of layers gave different results. Also, the time to train the model was different on the same number of training times. Specifically, we trained the ResNet models 20 times, the more layers the ResNet architecture had, the more time the ResNet model consumed. For example, ResNet-18 spent about 3 hours but ResNet-101 spent about 11.3 hours. Although consuming much time, the more layers ResNet architecture had, the higher the accuracy ResNet got. For example, ResNet-18 had a train loss was 99.4%, but ResNet-101 was 99.7%. From Table 4, we concluded that the more layers ResNet had, the more effective it was. The training time was slow but the accuracy was higher. Because the system only experimented with 8,375 images, the use of ResNet-34 architecture was extremely appropriate.

3.5. Accuracy comparison of restaurant recommendation with alternating least squares (ALS) and k-nearest neighbors algorithm (K-NN)

In this study, we used Spark to process and extract data about restaurant locations up to five million records. However, the structure of the restaurant information file had four columns such as food name, restaurant name, restaurant location, and rating while the ALS model accepted three columns. Therefore, we had to merge the restaurant name and restaurant address into one column named "Title" as shown in Table 5. Another issue was that the ALS model only accepted number values for all columns. So, we converted the food name as "Tag", and "Title" into two new columns "Tag_new", and "Title_new" which contained unique number values as shown in Table 6. Next, we split the restaurant information into two parts: 80% for training, and 20% for testing. Then, we applied the ALS model to predict the suitable restaurant location. The result of the recommendation prediction is shown in Table 7.

Table 5. Merging restaurant name and restaurant address into one column "Title"

Tag	Rating	Title
Banh beo	6.4	Quan Trang – Mi Quang
Banh beo	6.1	He thong banh beo Hue
Banh beo	6.5	Banh beo Hue Ngoc
Banh beo	7.6	Banh beo Hue – Binh Thanh
Banh beo	7.9	Banh beo Sai Gon

Table 6. Converting columns "Tag", "Title" which contained string into new columns "Tag_new"

Tag	Rating	Title	Title_new	Tag_new
Banh beo	6.4	Quan Trang – Mi Quang	22.0	0.0
Banh beo	6.1	He thong banh beo Hue	16.0	0.0
Banh beo	6.5	Banh beo Hue Ngoc	4.0	0.0
Banh beo	7.6	Banh beo Hue – Binh Thanh	2.0	0.0

Table 7. The result of recommendation prediction from the ALS algorithm

Tag	Rating	Title	Title_new	Tag_new	Prediction
Bun dau mam tom	7.7	Quan no – Bun dau	317.0	15.0	7.6994123
Goi cuon	8.0	Goi cuon tom nhay	212.0	24.0	7.9995100
Bun bo Hue	7.3	Me oi – Bun Bo Hue	289.0	14.0	7.298276

We implemented the K-NN algorithm to compare with ALS [27]. We randomly chose 10 food categories and used MAE and root RMSE to evaluate how well our recommendation was. The result of this comparison is shown in Table 8. In Table 8, the ALS algorithm achieved MAE and RMSE scores lower than K-NN. This pointed out that the performance of the recommender system implemented with ALS was better than the K-NN method.

Table 8. The comparison between ALS and K-NN on 10 food categories with two metrics MAE and RMSE

Food Name	ALS	K-NN
Bun dau mam tom	7.6	6.9
Goi cuon	7.9	5.7
Bun bo Hue	7.2	6.5
Bun thit nuong	7.6	7.7
Banh canh	6.8	6.8
Banh pia	8.5	7.2
Banh bot loc	8.4	7.4
Com tam	6.3	6.6
Banh mi	6.6	6.2
Banh beo	6.1	6.4
MAE	**0.1**	**0.7**
RMSE	**0.1**	**1.0**

3.6. Comparison of big data handling time

Growing information on the Internet rapidly, handling big data is an important task. We compared the speed of query processing between Spark and MS SQL Server. We split our restaurant dataset into three small datasets with 1 million, 2 million, and 5 million, respectively. From Table 9 and Figure 4, the comparison result showed that Apache Spark handled big data better than SQL Server in both data volume and processing speed.

Table 9. The query processing speed comparison between SQL-Server and SPARK

Data	1M	2M	5M
SQL-Server	10 (s)	16 (s)	46 (s)
Apache SPARK	5.3 (s)	8 (s)	40 (s)

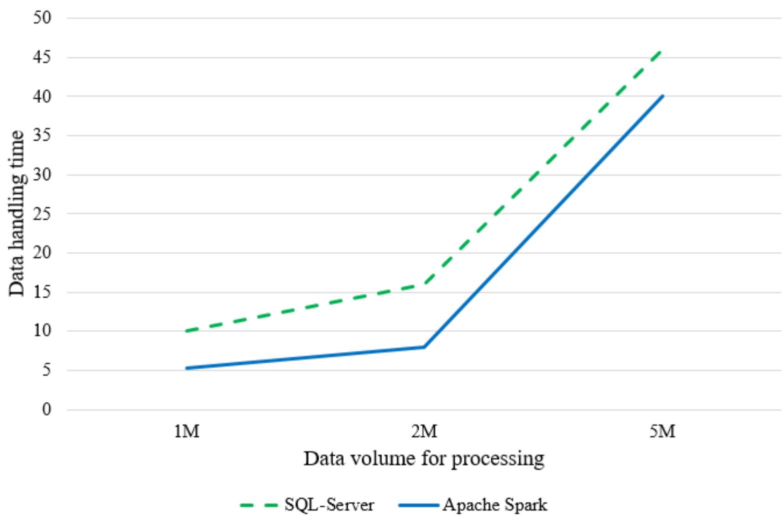

Figure 4. Query processing speed comparison between SQL-server and SPARK

4. CONCLUSION

The recommendation system has brought convenience to users in many fields such as e-commerce, music, and education, and has achieved significant results. However, most of the current recommender systems in the world in general and Vietnam in particular have also focused on developing in the field of e-commerce but have not yet been applied much to other fields. In this paper, we built a recommendation system in the tourism sector by combining machine learning and deep learning techniques. First, we used the ResNet-34 model to identify an input image to which class the image belongs. Next, we used the ALS method of matrix factorization to recommend the top 10 restaurant locations to tourists. In addition, the use of distributed computing in data processing tasks instead of traditional simple computations brought the output faster, and more accurately. Therefore, we found that this combination has achieved outstanding results. However, some limitations still exist. In this study, we only trained on the Vietnamese dataset of about 8,375 images and 5 million lines of recommended location data. These datasets are quite small. In addition, the data processing speed of the machine learning model has not been fully optimized because of the limitations of the machine resource as well as the volume of the dataset. Therefore, in future work, we could augment the dataset and use distributed computing on many computers to improve performance.

ACKNOWLEDGEMENTS

This study receives funding from Van Lang University, Ho Chi Minh City, Vietnam.

REFERENCES

[1] P. Do, T. H. V. Phan, and B. B. Gupta, "Developing a Vietnamese tourism question answering system using knowledge graph and deep learning," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 20, no. 5, pp. 1–18, Jun. 2021, doi: 10.1145/3453651.

[2] P. Do and T. H. V. Phan, "Developing a BERT based triple classification model using knowledge graph embedding for question answering system," *Applied Intelligence*, vol. 52, no. 1, pp. 636–651, May 2022, doi: 10.1007/s10489-021-02460-w.

[3] T. H. V. Phan and P. Do, "NER2QUES: combining named entity recognition and sequence to sequence to automatically generating Vietnamese questions," *Neural Computing and Applications*, vol. 34, no. 2, pp. 1593–1612, Sep. 2022, doi: 10.1007/s00521-021-06477-7.

[4] M. Hong and J. J. Jung, "Multi-criteria tensor model for tourism recommender systems," *Expert Systems with Applications*, vol. 170, Art. no. 114537, May 2021, doi: 10.1016/j.eswa.2020.114537.

[5] J. Sharma, K. Sharma, K. Garg, and A. K. Sharma, "Product recommendation system a comprehensive review," *IOP Conference Series: Materials Science and Engineering*, vol. 1022, no. 1, p. 12021, Jan. 2021, doi: 10.1088/1757-899X/1022/1/012021.

[6] F. Z. Trabelsi, A. Khtira, and B. El Asri, "Hybrid recommendation systems: A state of art," in *International Conference on Evaluation of Novel Approaches to Software Engineering, ENASE - Proceedings*, 2021, vol. 2021-April, pp. 281–288, doi: 10.5220/0010452202810288.

[7] J. Verbraeken, M. Wolting, J. Katzy, J. Kloppenburg, T. Verbelen, and J. S. Rellermeyer, "A survey on distributed machine learning," *ACM Computing Surveys*, vol. 53, no. 2, pp. 1–33, Mar. 2020, doi: 10.1145/3377454.

[8] M. S. Kurz, "Distributed double machine learning with a serverless architecture," in *ICPE 2021 - Companion of the ACM/SPEC International Conference on Performance Engineering*, Apr. 2021, pp. 27–33, doi: 10.1145/3447545.3451181.

[9] Z. Fayyaz, M. Ebrahimian, D. Nawara, A. Ibrahim, and R. Kashef, "Recommendation systems: Algorithms, challenges, metrics, and business opportunities," *Applied Sciences (Switzerland)*, vol. 10, no. 21, pp. 1–20, Nov. 2020, doi: 10.3390/app10217748.

[10] M. Vahidi Farashah, A. Etebarian, R. Azmi, and R. Ebrahimzadeh Dastjerdi, "A hybrid recommender system based-on link prediction for movie baskets analysis," *Journal of Big Data*, vol. 8, no. 1, Feb. 2021, doi: 10.1186/s40537-021-00422-0.

[11] M. Al-Ghamdi, H. Elazhary, and A. Mojahed, "Evaluation of collaborative filtering for recommender systems," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 3, pp. 559–564, 2021, doi: 10.14569/IJACSA.2021.0120367.

[12] A. Ouatiq, K. El-Guemmat, K. Mansouri, and M. Qbadou, "A design of a multi-agent recommendation system using ontologies and rule-based reasoning: Pandemic context," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 1, pp. 515–523, Feb. 2022, doi: 10.11591/ijece.v12i1.pp515-523.

[13] F. T. A. Hussien, A. M. S. Rahma, and H. B. Abdulwahab, "An e-commerce recommendation system based on dynamic analysis of customer behavior," *Sustainability (Switzerland)*, vol. 13, no. 19, p. 10786, Sep. 2021, doi: 10.3390/su131910786.

[14] N. Mauro, Z. F. Hu, and L. Ardissono, "Service-aware personalized item recommendation," *IEEE Access*, vol. 10, pp. 26715–26729, 2022, doi: 10.1109/ACCESS.2022.3157442.

[15] E. Kannout, "Context Clustering-based recommender systems," in *Proceedings of the 2020 Federated Conference on Computer Science and Information Systems, FedCSIS 2020*, Sep. 2020, pp. 85–91, doi: 10.15439/2020F54.

[16] A. Tripathi and A. K. Sharma, "Recommending restaurants: A collaborative filtering approach," in *ICRITO 2020 - IEEE 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)*, Jun. 2020, pp. 1165–1169, doi: 10.1109/ICRITO48877.2020.9197946.

[17] L. N. Trung, "Vietnamese_food," Kaggle. https://www.kaggle.com/datasets/lenguyentrung/vietnamese-food (accessed Oct. 05, 2023).

[18] L. N. Trung, "Vietnamese foods dataset," Kaggle. https://www.kaggle.com/datasets/nguyentrongdai/vietnamese-foods-dataset (accessed Oct. 05, 2023).

[19] Foody, "Places to eat, cuisine, where to eat in the city. HCM, restaurant, café," Foody.vn. https://www.foody.vn/ (accessed Oct. 05, 2023).

[20] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, no. 1, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

[21] A. A. Abdulmunem, N. D. Al-Shakarchy, and M. S. Safoq, "Deep learning based masked face recognition in the era of the COVID-19 pandemic," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 2, pp. 1550–1559, Apr. 2023, doi: 10.11591/ijece.v13i2.pp1550-1559.

[22] J. Liang, "Image classification based on RESNET," *Journal of Physics: Conference Series*, vol. 1634, no. 1, Art. no. 12110, Sep. 2020, doi: 10.1088/1742-6596/1634/1/012110.

[23] P. S. Rao, T. V. M. Rao, S. Kurumalla, and B. Prakash, "Matrix factorization based recommendation system using hybrid optimization technique," *EAI Endorsed Transactions on Energy Web*, vol. 8, no. 35, pp. 1–7, Jul. 2021, doi: 10.4108/eai.19-2-2021.168725.

[24] Y. K. Guntupalli, V. S. Saketh, S. Amudheswaran, and D. S. Vaishnav, "High-scale food recommendation built on Apache Spark using alternating least squares," Dec. 2020, doi: 10.1109/ICRAIE51050.2020.9358277.

[25] Z. A. Khan, N. I. Chaudhary, and S. Zubair, "Fractional stochastic gradient descent for recommender systems," *Electronic Markets*, vol. 29, no. 2, pp. 275–285, May 2019, doi: 10.1007/s12525-018-0297-2.

[26] S. T. Nguyen, H. Y. Kwak, S. Y. Lee, and G. Y. Gim, "Featured hybrid recommendation system using stochastic gradient descent," *International Journal of Networked and Distributed Computing*, vol. 9, no. 1, pp. 25–32, 2021, doi: 10.2991/ijndc.k.201218.004.

[27] J. He, G. Neubig, and T. Berg-Kirkpatrick, "Efficient nearest neighbor language models," in *EMNLP 2021 - 2021 Conference on Empirical Methods in Natural Language Processing, Proceedings*, 2021, pp. 5703–5714, doi: 10.18653/v1/2021.emnlp-main.461.

BIOGRAPHIES OF AUTHORS

Viet Hoang Pham received the B.Sc. degree in information technology from Van Lang University, Ho Chi Minh City, Vietnam, in 2022. Currently, he is a junior developer at DXC corporation. His research interests include computer vision and recommendation systems. He can be contacted at email: viet.187it14146@vanlanguni.vn.

Anh Thai Nguyen holds a B.Sc. in information technology from Van Lang University, Ho Chi Minh City, Vietnam, in 2022. He is currently a student majoring in data science and artificial intelligence at the Asian Institute of Technology, Thailand. His research interests include artificial intelligence, machine learning, and deep learning. He can be contacted at email: anh.187it14044@vanlanguni.vn or st122910@ait.asia.

Bao The Phung received the Ph.D. degrees in information technology from Irkutsk National Research Technical University (IrNITU), Irkutsk City, Russia, in 2014. He is currently the lecturer of Faculty of Information Technology, Van Lang University, Ho Chi Minh City, Vietnam. His research interests include the mathematical model, applications of artificial intelligence, deep learning, machine learning, data mining. He can be contacted at email: bao.pt@vlu.edu.vn.

Truong Ho-Viet Phan received the M.Sc. degrees in computer science from University of Information Technology, Ho Chi Minh Nation University, Vietnam, in 2013. He is currently the lecturer of Faculty of Information Technology, Van Lang University, Ho Chi Minh City, Vietnam. His research interests include the applications of artificial intelligence, deep learning, machine learning, text mining, social network analysis and chatbot. He can be contacted at email: truong.phv@vlu.edu.vn.