Analysis of entry behavior of students on job boards in Japan based on factorization machine considering the interaction among features

Tomoya Sugisaki\textsuperscript{1}, Yuri Nishio\textsuperscript{2}, Kenta Mikawa\textsuperscript{2**}, Masayuki Goto\textsuperscript{3} and Takashi Sakurai\textsuperscript{3}

Abstract: Job-hunting activities in Japan are different from those in other countries. The features of this are the simultaneous recruitment of new graduates, joining the company in April, and the use by most students of such resources as employment information websites. In recent years, website job boards for new graduates have provided Japanese students with assistance in finding companies for which they want to work. On these boards, students can bookmark companies that they are interested in before deciding to apply. After bookmarking, a company bookmarked by a user can examine the information again later. However, even if the students rate various companies, many of these bookmarks do not lead to job applications. In other words, this can be regarded as a lost opportunity for gaining job applications from the perspective of the company. It is important for companies to gain as many job applications as possible to be successful in their recruitment activities. Therefore, a method of analyzing the

ABOUT THE AUTHOR

Tomoya Sugisaki is currently a master’s candidate in industrial and management systems engineering from Waseda University, Tokyo, Japan. His research interest is machine learning. Yuri Nishio is currently a master’s candidate in industrial and management systems engineering from Waseda University, Tokyo, Japan. Her research interest is machine learning. Kenta Mikawa is an associate professor at Shonan Institute of Technology, Japan. He received his Doctor of Engineering degree from Waseda University in 2016. His research interests include machine learning and pattern recognition. Masayuki Goto is a professor at the Department of Industrial and Management Systems Engineering, School of Creative Science and Engineering, Waseda University, Tokyo, Japan. He received his Dr. of Engineering degree in Industrial Engineering from Waseda University in 2000. His research interests include how to effectively customize artificial intelligence models and machine learning algorithms in business domains. Takashi Sakurai is currently working at Recruit Career Co., Ltd. in data analytics related to a portal site for job-hunting.

PUBLIC INTEREST STATEMENT

Job-hunting activities in Japan are different from those in other countries. The features of this are the simultaneous recruitment of new graduates, joining the company in April, and the use by most students of such resources as employment information websites. In recent years, website job boards for new graduates have provided Japanese students with assistance in finding companies for which they want to work. A method of analyzing the entry behavior of students on job boards using factorization machines is proposed. The model predicts whether a student will submit a job application to a company. The prediction is based on student attributes and activity information, as well as information about the companies that they are interested in, as input variables. The interactions between input variables are also considered in making the prediction. In addition, the method supports student job-hunting activities and company measures for targeting students.
entry behavior of students on job boards using factorization machines is proposed. The model predicts whether a student will submit a job application to a company. The prediction is based on student attributes and activity information, as well as information about the companies that they are interested in, as input variables. The interactions between input variables are also considered in making the prediction. In addition, the method supports student job-hunting activities and company measures for targeting students. To clarify the proposed model, analytical experiments were conducted with actual data from a website job board for new graduates.

Subjects: Statistics for Business, Finance & Economics; Cybernetics; Machine Learning - Design

Keywords: Big data; management information; prediction; factorization machines
Subjects: Systems & Control Engineering; Machine Learning; Marketing; Statistics & Probability

1. Introduction

In recent years, the use of job boards on Japanese graduate websites for company recruitment and student job-hunting activities has been increasing. Companies can post information about recruitment, briefing sessions, internships, and job applications on the company page of the job board. Meanwhile, student users (hereinafter “users”) can participate in briefing sessions and apply for internships and jobs based on the information supplied. This means that there is a large amount of user attribute and activity information, as well as company information, on the job board. It is expected that such data will be used for various measures to support recruiting activity.

Some job boards have a function called on “interested list” where users can register companies in which they are interested. This makes it possible for them to receive employment information about registered companies through e-mails and so on. In short, these are bookmarks. In job-hunting activities, users need to choose specific companies that they are interested in working for from among a large number of companies. For this reason, many users register companies of interest in their bookmarks to facilitate their job-hunting activities. However, despite being bookmarked by users who are interested in them, many companies have not been able to translate these bookmarks into job applications. In other words, these can be seen as lost job application opportunities for many companies. If the job board webmaster can take appropriate measures for users with bookmarks, there is a high possibility that the number of job applications can be improved efficiently. Thus, it is relevant to predict whether a user will make job applications to a company if the company is registered in his/her bookmarks.

Therefore, in this study, a method is proposed for analyzing student entry behavior on job boards using factorization machines (FMs) (Rendle, 2010) by considering user attributes that include bookmarks and activity information, as well as company information, as input variables. The FM is a model that considers the interaction between input variables for prediction. The interaction in FM is represented by an inner product of vectors with relatively few parameters, so a relatively accurate estimate of interaction effects is possible in this model. Although there have been studies that used FMs for behavioral analysis (C. Chen et al., 2016; Wang et al., 2016), there have been no studies in which FMs were applied to such data as bookmarks. Clarifying the effects of interactions can support the decision making of users in job hunting and company targeting strategies. For example, when the model clarifies the positive effect of the interaction between “participation in a briefing session” (user activity information) and “humanities” (user attribute information) on the job applications for a company, the company can expect an increase in job applications from users by promoting briefing sessions for those studying in the humanities field.
Users who use the bookmark function on the job board are considered to be interested in registered companies, but many do not apply for jobs with the companies. Decreasing these lost opportunities is an important issue for companies. In this study, a binary classifier was constructed based on FM to predict the presence or absence of user job applications to a company utilizing user attributes, activity information, and company information as input variables. In addition, measures were considered for increasing the number of job applications by analyzing the interactions between the obtained features. By applying the FM classifier to actual data from a major website job board to confirm the accuracy of the binary classification, it was possible to investigate the possibility of classification and to identify important interactions. Furthermore, effective measures for increasing job applications from the relationship between features by analyzing secondary parameters obtained by the classification were investigated.

2. Job board on website in Japan and data for analysis

2.1. Overview
According to “Job Hunting Guide for International Students” (Japan Student Services Organization (JASSO), 2021), job-hunting activities in Japan differ from those in other countries. The features of this are the simultaneous recruitment of new graduates, who join the company in April. In addition, most students use such resources as employment information websites. Because of this, students are required to look for a job using employment information websites while they are still studying at a university.

A website job board (an Internet job-hunting portal) for new graduates is a web service that supports both company recruitment and user job-hunting activities. Companies can post recruitment information, such as briefing sessions, internships, and job descriptions, as well as basic information on the company page on the job board. The purpose is to solicit job applications from users (new graduates) on the board. Users, however, can participate in briefing sessions and apply for internships with companies in which they are interested, based on the information that the companies publicize, and then apply for jobs with them. In recent years, many companies have published their information on boards, so users can select prospective employers from a wide range of information. From this perspective, the purpose of the bookmark function provided on the job boards is to facilitate the experience and job search of users. Users can bookmark multiple companies that they are interested in and easily receive recruiting information from the companies through e-mail.

When using the job board service, it is necessary for companies to register with the site management company and request to post information. Similarly, the user must register an account on the job board, and basic information, such as university name, department, and subject majors, should be registered. Therefore, the job board not only has user activity information (such as participation in briefing sessions, internships, and job applications) but also basic user information. Examples of the main accumulated data are shown in Table 1.

When users log in to the job board, they can search by business type or keyword for companies that they are interested in or gather information by following site recommendations. In addition,
they can bookmark the companies in which they are interested and check the information later. Subsequently, they can deepen their understanding through participating in briefing sessions or internships and, finally, apply for jobs with the companies for which they want to work.

### 2.2. Previous research on job-hunting and recruiting activities

Job-hunting and recruitment activities have already been the subject of much Japanese research. For example, Shimomura and Hori (2004) and Nagano (2005) conducted sociological studies based on interviews and field surveys without using data accumulated on website job boards. However, various studies have been conducted on the relationship between company appeal and the reasons why users apply (Sakamoto et al., 2016), models for analyzing the relationship between student job board browsing and actually making job applications (Sugiyama et al., 2017), and models for predicting the number of job applications (Nagamori et al., 2016; Nodu et al., 2015) using data on the job board. However, the interaction of the variables contributing to the number of job applications has not been examined in these prediction models. Considering these interactions makes it possible not only to improve the prediction accuracy, but also to narrow the user target layer and take measures to increase the number of job applications. The major difference between this study and the above-mentioned studies is the topic of whether to target users who use the bookmarking function. Using the bookmarking function, one can extract the features of the companies in which students are interested. The aim of this study is to reduce the lost opportunities for companies by using information from the bookmarks. Therefore, a study focusing on the interaction between variables is valuable.

### 2.3. Brief information about bookmarks

In recent years, many companies have posted their information on graduate job board websites, so users have a wide range of companies from which to select. However, it is difficult for users to peruse large amounts of recruitment information from a large range of interesting companies one by one on multiple occasions; consequently, they may overlook a desired company. Therefore, the job boards introduced the bookmark function to solve the problem. Users can bookmark companies that they are interested in and can easily receive recruiting information through e-mail or find company information without searching. They can collect information on companies of interest efficiently, and they can easily proceed with job-hunting activities by utilizing this list. Approximately 74% of users are using the bookmark function, showing that it plays an important role in job-hunting activities. It can be assumed that a user is interested in a company when he/she bookmarks the company. Therefore, it is considered that a user who bookmarks a company is more likely to apply for a job with that company than a user who does not bookmark it.

The ratios of users who are taking actions (job applications, briefing reservations, and internship reservations) with bookmarked companies out of the users who are using the bookmark function is approximately 58%. However, fewer users engage in briefing sessions or internship reservations than apply for jobs. From this result, it appears that there are many users who have not taken

| Types of data     | Item name                                                                 |
|-------------------|---------------------------------------------------------------------------|
| Attribute information | User ID, school name, department, humanities or science, etc.             |
| Activity information | Briefing reservation, internship briefing reservation, internship reservation, job application, etc. |
| Company information | Company ID, company name, business types, number of employees, presence or absence of stock offering, etc. |
action on any company despite their initial interest. It appears that companies have lost the opportunity to gain job applications from users who were temporarily interested. Therefore, they are challenged to take measures to promote themselves to those users.

2.4. Summary of dataset

The aim of this study was to analyze users who utilize the bookmark function on the job board. Here, an overview of the dataset is provided. It was collected by a Japanese company (referred to as “company A”) managing a job board site (a portal web site for job hunting) on the Internet during the period from June 2015 to March 2017 to analyze the entry behavior of students. The data to be analyzed are shown in Table 2.

The features used are shown in Table 3. In addition, the numbers in () represent the number of types of feature.

Concrete and brief explanations about the features in Table 3 are shown as follows;

- “Affiliation”: It has 2 features that the student belongs to “humanity department” or “science department”. In Japan, the categories of “humanity department” and “science department” at universities are commonly used.
- “Old address”: The prefecture the user used to live in. It has 48 features because this is a 1-hot vector representing 47 prefectures in Japan adding overseas.
- “Current address”: The prefecture the user is living in now. It has 48 features same as the current address.
- “Learning classification”: It has 2 features of “graduate student” or “undergraduate student”.
- “Briefing reservation”: It means whether the student reserved a briefing session or not.
- “Internship briefing reservation”: It means whether the student reserved an Internship briefing session or not.
- “Application for internship”: It means whether the student applied for an internship or not.
- “Business type”: The type of business that the company engaged. It has 126 features because this is a 1-hot vector representing 126 types of the business divided by a website job board.
- “Stock offering”: It means whether the company presents or absents of stock offering. It has 2 features because this is a 1-hot vector representing the presence or absence of that.
- “Company size”: The company size that is divided into 8 discrete levels.
- “Head office location”: The prefecture where the company’s head office location is. It has 47 features because this is a 1-hot vector representing 47 prefectures in Japan.

Note that the sum of all features is $I = 286$.

| Table 2. Data to be analyzed | From June 2015 to March 2017 |
|-----------------------------|-----------------------------|
| **Period covered**          |                             |
| Target user condition       |                             |
|                             | The number of bookmarks is from 6 to 9 |
|                             | University student or graduate student |
| **Number of features**      | $I = 286$                   |
| **Number of learning data** | 119,040                      |
| **Number of test data**     | 29,760                       |
Table 3. Features used this analysis

| Attribute information | Name (the number of features) | Meaning                        |
|-----------------------|-------------------------------|--------------------------------|
| Old address (48)      | Prefectures or overseas       |                                |
| Current address (48)  | Prefectures or overseas       |                                |
| Learning classification (2) | University or graduate student |                                |

| Activity information | Name (the number of features) | Meaning                        |
|----------------------|-------------------------------|--------------------------------|
| Briefing reservation (1) | Reserve a briefing session     |                                |
| Internship briefing reservation (1) | Reserve an internship briefing session |                                |
| Application for internship (1) | Application for internship     |                                |

| Company information | Name (the number of features) | Meaning                        |
|---------------------|-------------------------------|--------------------------------|
| Business type (126) | Business type                 |                                |
| Stock offering (2)  | Presence or absence of stock offering |                                |
| Company size (8)    | Company size divided into 8 stages |                                |
| Head office location (47) | Prefecture                  |                                |

3. Factorization machines

3.1. Overview
The FM model considers the interaction between features of data and is known to exhibit high prediction accuracy. In addition, it has a high generalization performance and has been applied to various problems (Blondel et al., 2016; Rendle, 2012). However, the majority of studies on FM have focused on improving the accuracy, and there are only a few studies in which knowledge discovery was performed with the help of the obtained models (C. Chen et al., 2016; Wang et al., 2016). Furthermore, in these studies, the obtained interaction terms were not analyzed.

Generally, if the number of features of input data is I, then the number of parameters of two-way interactions is in proportion to I². Therefore, as I increases, the number of data required for parameter estimation increases enormously. Considering the above, FM expresses the interaction between features with a relatively small number of parameters by calculating the inner product of each row of a low-dimensional matrix called the interaction matrix (I x K(K ≪ I)).

3.2. Formulation
Consider N pairs of n-th explanatory variable vectors \(x_n = (x_{n1}, x_{n2}, \ldots, x_{nd})^T\), \((x_{ni} \in \{0,1\}, i = 1,2,\ldots,I)\) and objective variable \(y_n \in \mathbb{R}\). Let \(w_0\) be the bias term and \(w = (w_1, w_2, \ldots, w_I)^T\) be the weight vector. The interaction matrix is represented as a matrix \(V = [v_1, v_2, \ldots, v_I]^T \in \mathbb{R}^{I \times K}\) with the K dimensional vector \(v_n = (v_{1n}, v_{2n}, \ldots, v_{Kn})^T\) as an element.

As shown in Figure 1, the input data are a vector having 1 when each feature amount is included, and 0 when it is not. If the variables defined above are used, the FM model for \(x_n\) is given by:

\[
f(x_n) = w_0 + \sum_{i=1}^{I} w_i x_{ni} + \sum_{i=1}^{I} \sum_{j=i+1}^{I} v_{ij} x_{ni} x_{nj}
\] (1)
The right-hand side of Equation (1), when only the first and second terms are used, is the same as the expression of the multiple regression model. In FMs, the prediction accuracy is improved by adding a third term representing the interaction between features. This is represented by Equation (2):

$$v_i, v_j = \sum_{k=1}^{K} v_{ik} \cdot v_{jk}$$  \hspace{1cm} (2)

### 3.3. Characteristics of FMs

Unlike the general regression model, an FM makes it possible to improve prediction accuracy by considering the interaction term of features according to the third term of Equation (1). As shown in Equation (3), by calculating the inner product of each row of a low-dimensional matrix $V \in \mathbb{R}^{I \times K}$, called an “interaction matrix,” the number of parameters decreases from $I^2$ to $I \times K (I \ll K)$, and the interaction can be represented by an inner product of vectors with relatively few parameters.

$$V = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1K} \\ v_{21} & v_{22} & \cdots & v_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ v_{I1} & v_{I2} & \cdots & v_{IK} \end{pmatrix}$$ \hspace{1cm} (3)

As a result, the number of parameters does not increase excessively, and it can reduce the degrees of freedom and avoid overfitting.

### 3.4. Extension method of the FM and its application

An FM is an effective prediction model; accordingly, several improved FMs for actual applications to solve real-world problems have been reported. Y. Chen et al. (2019) extended FMs for personal feature interaction selection to improve recommendation performance using Bayesian variable selection. In real-world problems, input data often contain noise. To solve this problem and achieve robust parameter estimation, robust parameter estimation methods using FMs have been proposed. Punjabi and Bhatt (2018) applied an FM for user response prediction, a task that often involves noisy data. To achieve accurate prediction, they proposed a robust factorization machine under a robust optimization scheme. Moreover, Ni et al. (2018) extended an FM to class imbalance and noisiness settings to apply the FM to credit card default prediction. They defined a nonconvex loss function (asymmetric ramp loss) to the FM and demonstrated the robustness and effectiveness of the proposed method.

However, some studies have used FMs for behavioral analysis. Wang et al. (2016) applied an FM to analyze the behavior of customers on e-commerce sites to predict the behavior of customers. In addition, C. Chen et al. (2016) used FMs with behavioral analysis to predict product purchases. Using feature engineering, they showed that accurate predictions can be made. These studies used FMs for behavioral prediction, but they focused only on prediction accuracy and did not attempt to take advantage of the estimated interactions.

### 4. Application of FM for classification

In this article, a job application behavior analysis model is proposed to increase the number of job applications. The proposed model predicts whether a user will apply for a job with a company registered in his/her bookmarks. The model employs the attributes, activity information, and bookmarked company information of the user as input variables. Analyzing the interaction makes it possible to analyze the relationship between the features that influence job applications.

However, the FM described in the previous section is known as a prediction model that considers the interaction between features. In research reported here, an FM was applied to the prediction of whether a user will apply for a job with a company registered in his/her bookmarks. Furthermore, the effective relationships between features are analyzed from the obtained parameters. For general binary classification, with 0.5 as the boundary, a value of 0.5 or more was set to be 1 and of less to be 0. However, the FM represented by Equation (1) is a general regression model.
(hereinafter “regression FM”), and the output is a real value. Therefore, to obtain a model suitable for binary classification, the output obtained by the regression FM is taken as the input of the logistic function (Bishop, 2007), and an output value of (0, 1) is obtained.

### 4.1. Formulation of logistic FM

To modify the regression FM to the binary classification, the output of the regression FM expressed by Equation (1) is input to the logistic function. A value of 0.5 or more was set to be 1 and less to be 0 with 0.5 as the boundary. At this time, the model equation is expressed as Equation (4), which is called the “logistic FM”:

$$g(x_n) = \frac{1}{1 + \exp(-f(x_n))}$$  \hspace{1cm} (4)

When classifying a new datum \(x\), it is considered 1 if \(g(x)\) is greater than 0.5, and 0 if it is less than 0.5. The number of explanatory variables is 286, as shown in Table 2.

The difference between logistic regression and this logistic FM is whether the interactions between the features are or are not considered. The method of parameter estimation of logistic FM is also based on the gradient method.

### 4.2. Parameter estimation

When the objective variable is \(y_n \in \{0, 1\}\), the objective function is considered to maximize the likelihood function (Myung, 2003) to obtain the optimal parameters for the logistic FM. The likelihood function for \(N\) numbers of training data is expressed as follows:

$$L(w_0, w, V) = \prod_{n=1}^{N} g(x_n)^{y_n} \{1 - g(x_n)\}^{(1-y_n)}$$  \hspace{1cm} (5)

Here, maximizing the likelihood function is equivalent to minimizing the negative log-likelihood function. Considering a negative log-likelihood function, the objective function of the parameter estimation can be given by Equation (6):

$$LL(w_0, w, V) = -\log L(w_0, w, V) = \sum_{n=1}^{N} \{y_n \log g(x_n) + (1 - y_n) \log(1 - g(x_n))\}$$  \hspace{1cm} (6)

Based on the above discussion, the parameter estimation minimizing the negative log-likelihood function \(LL\) expressed by Equation (6) is considered. Furthermore, “regularizer” terms with regularization parameters \(\lambda\) and with \(\ell_2\) the norm of vector (Hoerl & Kennard, 2000) are introduced to prevent overfitting. Here, the objective function is expressed by Equation (7):

$$\min \sum_{n=1}^{N} \{y_n \log g(x_n) + (1 - y_n) \log(1 - g(x_n))\} + \frac{1}{2} \lambda \|w\|_2^2 + \frac{1}{2} \lambda \sum_{j=1}^{J} \|v_j\|_2^2$$  \hspace{1cm} (7)

| Table 4. Experimental result |
|--------------------------------|-----------------|-----------------|
|                                | Linear logistic regression | Logistic FM     |
| Accuracy                       | 0.664            | 0.687           |
| Recall                         | 0.685            | 0.660           |
| Precision                      | 0.683            | 0.710           |
| F-measure                      | 0.684            | 0.683           |
Here, let $\cdot_2$ denote the $\ell_2$ norm of vector. In this study, the stochastic gradient descent method (hereinafter “SGD”) is introduced (Boyd & Vandenberghe, 2004) when minimizing Equation (7). The update formula for each parameter is expressed by Equations (8)–(10). Here, $w_0^{\text{old}}, w_i^{\text{old}},$ and $v_i^{\text{old}}$ are parameters before the update, $w_0^{\text{new}}, w_i^{\text{new}},$ and $v_i^{\text{new}}$ are parameters after the update, and $\alpha$ is the learning rate.

\begin{equation}
    w_0^{\text{new}} := w_0^{\text{old}} - \alpha \{ g(x_n) - y_n \} 
\end{equation}

\begin{equation}
    w_i^{\text{new}} := w_i^{\text{old}} - \alpha \{ g(x_n) - y_n \} x_{ni} + \lambda w_i 
\end{equation}

\begin{equation}
    v_i^{\text{new}} := v_i^{\text{old}} - \alpha \left[ \{ g(x_n) - y_n \} x_{ni} \sum_{j=1}^{l} v_{jk} x_{nj} - v_{ik} x_{ni}^2 \right] + \lambda v_i 
\end{equation}

5. Data analysis
To show that the logistic FM is effective in binary classification problems, an accuracy evaluation was performed by using actual data accumulated on a job board for new graduates. Furthermore, the interactions were analyzed from the estimated parameters obtained as a result of learning, and effective measures were examined to increase the number of users making job applications.

5.1. Analysis condition
In this analysis, $y_n \in \{0, 1\}$ is the objective variable of the test data, which takes 1 for applying for a job and 0 for not applying. The number of explanatory variables is 286, as shown in Table 2. In addition, when a prediction is made, a case where the value obtained from Equation (4) is 0.5 or more is set as a job application, whereas less than 0.5 is taken as no job application. Accuracy, precision, recall, and F-measure (Manning et al., 2008) were used as the evaluation criteria in the experiment. In addition, the parameters were set based on preliminary experiments. The number of rows or parameter $K$ of the interaction matrix, the learning rate $\alpha$ of SGD, and the regularization parameter $\lambda$ were set to 13, 0.001, and 0.00008, respectively. There are no clear evaluation criteria for these parameters. Therefore, a preliminary experiment was performed to determine the values of these parameters.

| Table 5. Ranking of values of $w_i$ |
|--------------------------------------|
| Rank | Feature value    | $w_i$   |
|------|------------------|--------|
| 1    | Affiliation A    | +1.182 |
| 2    | Stock offering A| +1.085 |
| 3    | Affiliation B    | +0.715 |
| 4    | Head office location A | +0.708 |
| 5    | Old address A    | +0.551 |
| ...  | ...              | ...    |
| 282  | Company size A   | -0.360 |
| 283  | Business type A  | -0.363 |
| 284  | Business type B  | -0.412 |
| 285  | Business type C  | -0.430 |
| 286  | Business type D  | -0.535 |
5.2. Experimental result

First, to verify the performance of the FM binary classifier with the target problem, the prediction results were compared with the linear logistic regression, which is a general binary classification method. Fivefold cross validation (Bishop, 2007) was performed using the above data, and the results of each evaluation index for the test data in each method are shown in Table 4. In the table, a bold number indicates that it is better.

Table 4 shows that the logistic FM exceeded the correct answer rate by approximately 2.3% over linear logistic regression, and it was able to obtain almost the same value as the F-measure. From this result, one can see that the logistic FM obtains a higher accuracy than the linear logistic regression model and is a suitable model for predicting whether users will make job applications. In addition, the recall ratio is 34% and is the proportion of users who were predicted to make a job application and did not, in fact, do so. In other words, this 34% are the users who are judged by the model's prediction result to have a high possibility for

| Rank | Feature value     | Interaction effect |
|------|-------------------|--------------------|
| 1    | Current address A | +1.007             |
| 2    | Current address B | +0.930             |
| 3    | Business type E   | +0.887             |
| 4    | Business type A   | +0.807             |
| 5    | Business type F   | +0.805             |
| ...  | ...               | ...                |
| 282  | Business type G   | -0.910             |
| 283  | Business type H   | -0.920             |
| 284  | Business type I   | -1.040             |
| 285  | Business type J   | -1.126             |
| 286  | Business type K   | -1.467             |

Table 6. Interaction with “internship briefing reservation”

| Rank | Feature value     | Interaction effect |
|------|-------------------|--------------------|
| 1    | Current address A | +1.007             |
| 2    | Current address B | +0.930             |
| 3    | Business type E   | +0.887             |
| 4    | Business type A   | +0.807             |
| 5    | Business type F   | +0.805             |
| ...  | ...               | ...                |
| 282  | Business type G   | -0.910             |
| 283  | Business type H   | -0.920             |
| 284  | Business type I   | -1.040             |
| 285  | Business type J   | -1.126             |
| 286  | Business type K   | -1.467             |

Table 7. Interaction with “application for internship”

| Rank | Feature value     | Interaction effect |
|------|-------------------|--------------------|
| 1    | Business type A   | +1.175             |
| 2    | Head office location B | +1.122 |
| 3    | Business type E   | +1.089             |
| 4    | Business type M   | +1.049             |
| 5    | Business type F   | +1.048             |
| ...  | ...               | ...                |
| 282  | Business type H   | -1.065             |
| 283  | Business type L   | -1.065             |
| 284  | Business type J   | -1.163             |
| 285  | Business type I   | -1.266             |
| 286  | Business type K   | -1.412             |
making a job application. Therefore, an increase in the number of job applications can be expected by analyzing the causes and trends of the misclassifications and taking action with the users.

5.3. Analysis of weight vector \( w \)

The estimated weight vector \( w \) of the logistic FM was analyzed. The top-five and bottom-five values of \( w_i \) are shown in Table 5. The larger the value of \( w \), the more likely it is that a job application will be made as a result of that feature. Conversely, the smaller it is, the less likely it is to influence job applications.

From Table 5, the feature “Head office location A” appears in the top rank. This indicates that companies whose headquarters are located in A are more likely to receive job applications from users than other bookmarked companies. For this reason, it is important for companies with “head office location A” to take measures to be bookmarked. However, “business type D” is an example of a feature quantity appearing in the lower rank. In other words, companies of business type D tend to be bookmarked but are not linked to job applications made any more than companies of other business types. Therefore, business type D companies need to take appropriate measures after being bookmarked to increase the number of job applications they receive. In addition, for an affiliation that has the same kind of feature quantity, “affiliation A” has a larger value than “affiliation B,” so it can be interpreted that it is more likely for users with affiliation A to make a job application. As described above, by comparing various feature quantities, it is possible to extract features that are likely to lead to job applications.

5.4. Analysis of interaction between features

The relationship between features \( i \) and \( j \) is quantified based on the inner products of the obtained interaction vectors \( v_i \) and \( v_j \), and effective measures for increasing job applications are discussed with reference to these values. Here, the focus is on the interaction of each feature with “internship briefing reservation.” Table 6 shows the top-five and the bottom-five of the inner product between the internship briefing reservation and each other feature. It can be interpreted that, if the value of interactions is large, it tends to lead to a job application; conversely, if it is small, it does not. From Table 6, one can see that “current address A” and “current address B” appear in the upper features representing the effect of the interaction with the internship briefing reservation. In other words, it is more likely for a user of current address A or B to make a job application to a company when a reservation for an internship briefing was made with that company. This means that companies can expect an increase in the number of job applications by promoting participation in internship briefing sessions with users whose current addresses are in area A or B.

Furthermore, “business type A,” whose \( w_i \) was included in the lower-five ranking, is included in the upper ranking of the interactions with the internship briefing reservation. In other words, one can conclude that “business type A” is a company category that does not easily receive job applications as a whole. However, when users make appointments for internship briefing sessions, they tend to make job applications. Based on this result, companies of “business type A” can expect to increase the number of job applications by encouraging users to make appointments for internship briefings. In this way, by analyzing the interactions between various feature quantities, it is possible to identify users with feature quantities that are highly likely to lead to job applications. Therefore, the company can increase its success by promoting users who are highly likely to apply for jobs.

Table 7 shows the top five and bottom five of the inner product between an application for internship and each other feature. It is clear that “Head office location B” is newly included in the upper rank, unlike in Table 5. This shows that a company whose head office location might be related to apply the internship briefing for the user. The user application for the internship is also an important factor leading to a job application. Therefore, strongly encouraging the user
who applied for the internship to apply for a job will greatly affect the number of job applications.

6. Discussion

Through analysis using real data, it was shown that applying the logistic function to an FM is effective as a binary classification model. By analyzing secondarily obtained parameters, it was possible to find relationships between effective features in planning measures to increase the number of job applications to several companies.

Users who utilize website job boards for new graduates change every year. One can therefore infer that, because there is no identical user in the learning period and the prediction period, it is essentially as different a problem as the prediction of the purchasing behavior of each customer in the retail industry in this respect. However, even though the year changes, the basic trend of the user behavior does not change significantly. Based on this, the proposed model can consider not only user behavior information but also the tendency of user behavior as a feature quantity, ensuring even higher prediction accuracy. This can be realized by including behavior information that contains time series in the features.

By analyzing the interactions, it was possible to find a combination of feature quantities effective for ensuring job applications. By finding and analyzing the combination of feature quantities effective for each company, it is possible to develop measures to increase the number of job applications. In this study, only the interactions that focused on the internship briefing reservation feature were analyzed. However, by changing the feature focused on, various interactions can be analyzed, and the possibilities for increasing the number of job applications can be widened.

The FM model obtains higher accuracy than a general linear model. However, the accuracy of the correct answer rate was not improved significantly in this study. One reason for this result may be that the model is not effective if the data are too sparse. In general, the FM is considered to be effective with sparse data by considering a low-dimensional vector expressing interactions, but it is believed that there is a limit to the degree of data sparsity, which can be effectively modeled. Therefore, considering countermeasures against such overly sparse data can be a topic for future research.

7. Conclusion and future works

In this research, the logistic FM was applied to analyze the data of an Internet job board site for new graduates, and a job application behavior analysis model was constructed. By applying the model to analyze actual data, it was shown that the model is effective for planning measures to increase the number of job applications received by considering the strength of the interactions of the features. It was demonstrated that there are business groups that can take effective measures by considering the ranking of interactions.

As a future task, the time series of user behavior information should be considered. In this research, only the presence or absence of behavior was considered; however, if the order of actions changes, the meaning of each action changes. By considering the time series, it should be possible to express the behavior of users more accurately. In addition, it is necessary to check how many sparse data can be used.

Funding

The authors received no direct funding for this research.

Author details

Tomoya Sugisaki
E-mail: sugi5524@ruriu.waseda.jp
Yuri Nishio
Kenta Mikawa
E-mail: mikawa@info.shonan-it.ac.jp

Masayuki Goto
Takashi Sakurai
E-mail: tac_sakurai@r.recruit.co.jp

1 Department of Industrial and Management System Engineering, School of Creative Science and Engineering, Waseda University, Tokyo, Japan.
2 Department of Information Science, Faculty of Engineering, Shonan Institute of Technology, Fujisawa, Japan.
Sugisaki et al., Cogent Engineering (2021), 8: 1988381
https://doi.org/10.1080/23311916.2021.1988381

3 Recruit Career Co., Ltd, Tokyo, Japan.

Citation information
Cite this article as: Analysis of entry behavior of students on job boards in Japan based on factorization machine considering the interaction among features, Takayuki Sugisaki, Yuri Nishio, Kenta Mikawa, Masayuki Goto & Takashi Sakurai, Cogent Engineering (2021), 8: 1988381.

Notes
1. The originator of FM, S. Rendle stated that, when using FM for classification problems, logistic functions can be used. In this study, it is called a “logistic FM.”
2. Under the terms of the confidentiality agreement, it is not possible to indicate the name of the company. Therefore, the expressions A, B, etc. are used.

Disclosure statement
No potential conflict of interest was reported by the author(s).

References
Bishop, C. M. (2007). Pattern recognition and machine learning. Springer.
Blondel, M., Fujino, A., Ueda, N., & Ishihata, M. (2016). Higher-Order factorization machines. Proceedings of the 30th International Conference of Neural Information Processing Systems (pp. 3359–3367).
Boyd, S., & Vandenberghe, L. (2004). Convex optimization. Cambridge University Press.
Chen, C., Hou, C., Xiao, J., & Yuan, X. (2016). Purchase behavior prediction in E-Commerce with factorization machines. IEICE Transactions on Information and Systems, E99-D(1), 270–274. doi:10.1587/transinf.2015EDL8116
Chen, Y., Ren, P., Wang, Y., & Rijke, M. (2019). Bayesian personalized feature interaction selection for factorization machines. SIGIR’19: Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, Paris, France, 665–674. doi:10.1145/3331184.3331196
Hoel, A. E., & Kennard, R. W. (2000). Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 42(1), 80–86. doi:10.1080/00401706.2000.10485983

Japan Student Services Organization (JASSO). (2021). Job hunting guide for international students 2021. https://www.jasso.go.jp/eringyousa_icsFiles/fieldfile/2020/03/03/guide2021_all_e.pdf

Manning, C., Raghavan, P., & Schütze, H. (2008). Introduction to information retrieval. Cambridge University Press.

Myung, J. (2003). Tutorial on maximum likelihood estimation. Journal of Mathematical Psychology, 47(1), 90–100. doi:10.1016/S0022-2496(02)00028-7

Nagamori, S., Yamashita, H., Goto, M., & Ogihara, T. (2016). Proposal of prediction model for number of entries of job hunting portal site based on mixed regression. Proceedings of Symposium on Information Theory and its Application (ISITA2015), No. 7.3.4. (in Japanese).
Nagano, H. (2005). Employment awareness as a factor of job hunting success. The Review of Economics & Political Science, 73(5–6), 93–113. http://hdl.handle.net/10291/13483 (in Japanese).
Ni, W., Liu, T., Zeng, Q., Zhang, X., Duan, H., & Xie, N. (2018). Robust factorization machines for credit default prediction. Proceedings of the 15th Pacific Rim Conference on Artificial Intelligence, PRICAI 2018, Nanjing, China.
Nodu, T., Mikawa, K., Goto, M., & Ogihara, T. (2015). A Study on Prediction Model of the Number of Applications on Internet Portal Sites for Job Hunting. IEICE Technical Report, 115(381), 49–54. AI2015-34. (in Japanese).
Punjabi, S., & Bhatt, P. (2018). Robust factorization machines for user response prediction. WWW ’18: Proceedings of the 2018 World Wide Web Conference, Lyon, France, 669–678.
Rendle, S. (2010). Factorization machines. Proceedings of IEEE International Conference on Data Mining, Sydney, Australia, 995–1000.
Rendle, S. (2012). Factorization machines with libfm. ACM Transactions on Intelligent Systems and Technology, 3(3), 57–78. http://doi.acm.org/10.1145/2168752.2168771
Sakamoto, T., Yamashita, H., Goto, M., & Ogihara, T. (2016). A latent class model to analyze the relationship between companies’ appeal points and students’ reasons for application. Proceedings of the 7th Forum for Council of Industrial Engineering and Logistics Management Department Heads (CIEDH2016) & The 5th Institute of Industrial and Systems Engineering Asian Conference (IISEAsia2016), No. 109, Hong Kong.
Shimomura, H., & Hori, H. (2004). Technological exploratory behavior in job hunting for college students. Japanese Journal of Social Psychology, 2012, 93–105. https://doi.org/10.14966/jssp.KJ00003724981 (in Japanese).
Sugiyama, Y., Arai, T., Yang, T., Goto, M., & Ogihara, T. (2017). An analytical model of relation between browsing and entry activities on an internet portal site for job hunting. Proceedings of 15th Asian Network for Quality Conference (ANQ2017), Kathmandu, Nepal.
Wang, Y., Shang, W., & Li, Z. (2016). The application of factorization machines in user behavior prediction. 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), 1–4, Okayama, Japan.
