

## ORIGINAL RESEARCH

# Study of hiring decisions by companies using text mining: Factors other than experience

Hiroimi Asano\*, Koji Tanaka, Kazuhiko Tsuda

*Graduate School of Business Sciences, University of Tsukuba, Tokyo, Japan*

**Received:** May 14, 2016

**Accepted:** August 14, 2016

**Online Published:** August 22, 2016

**DOI:** 10.5430/air.v6n1p16

**URL:** <http://dx.doi.org/10.5430/air.v6n1p16>

## ABSTRACT

This study attempted to clarify the factors that lead to an applicant receiving a job offer, by applying text mining to feedback comment of notification form from recruiters. The analysis identified experience as the most salient factor. However, there were frequent cases in which the applicant was rejected despite the evaluation of experience being positive, or hired despite the evaluation of experience being negative. This paper focuses on the cases in which the result reversed the evaluation of experience. Our conclusions were as follows: 1) in many cases where experience was required, job offers were based on determinants other than experience, 2) key determinants that offset the evaluation of experience were motivation, aspiration, youth, and qualifications, 3) in jobs requiring experience, offers were made based on knowledge and skill as well as experience, whereas for jobs that did not require experience, offers were based on age and personality.

**Key Words:** Text mining, Hiring or rejection factors, Word extraction analysis, Job matching, Experience

## 1. INTRODUCTION

In recent years, employment conditions have been improving in Japan. Since November 2013, the effective job opening-to-application ratio, which is the monthly number of job openings listed by the Public Employment Security Office (Hello Work) divided by the effective number of applicants, has exceeded 1. The number of job openings exceeds the number of job seekers (Ministry of Health, Labour and Welfare).<sup>[1]</sup> Despite this apparent labor shortage, a large number of job applicants are still being rejected.

Companies decide whether to hire an employee based on the specific resources they seek. This is true in both Japan and European countries. However, as pointed out by Hamaguchi, the West has a “job-based labor society” in which an employee matching the occupation is hired, whereas Japan has a “membership-based labor society” in which the occupation

is matched to the employee.<sup>[2]</sup> The precise nature of the job is not necessarily clear in Japan, for companies of all sizes, excluding certain specialist professions. For this reason, it cannot be said that in Japan the decision to hire is based solely on the vocational skills of the applicant. Although the decision to hire still involves matching the vocational ability with labor requirements, other hiring or rejection factors are in play, and these are based on tacit knowledge. It is difficult for the job seeker to deduce the nature of these factors even after selection. In other cases, although the hiring company claims to reveal the vocational skills that were sought in the job seeker, the successful applicant does not appear to possess those skills. Even for the hiring company, the question of what is actually evaluated when determining to hire can be said to be tacit.

If the factors that make up this tacit knowledge can be clari-

\* **Correspondence:** Hiroimi Asano; Email: [asano@gssm.otsuka.tsukuba.ac.jp](mailto:asano@gssm.otsuka.tsukuba.ac.jp); Address: Graduate School of Business Sciences, University of Tsukuba, 3-29-1, Otsuka, Bunkyo-ku, Tokyo 112-0012, Japan.

fied, the efficiency of external labor markets in Japan ought to increase.

This study examined the factors that determine whether a candidate was hired or rejected by focusing on free comments from the information collected by job referral agencies after the job application process was complete. These are thought to reveal the true intentions of the hiring company. This study used the free comment data that is collected from the hiring company by Hello Work, Japan's largest labor supply and demand regulation agency. Applying text mining to this free comment data, the key factors in hiring or rejection to determine the factors that make up this tacit knowledge were analyzed.

## 2. PREVIOUS RESEARCH

Previous research on job selection in Japan is as follows.

The Ministry of Economy, Trade and Industry, Japan Business Federation, Japan Association of Corporate Executives, and others have conducted surveys on companies recruiting new graduates, and deduced that the most highly regarded skill are the roles played by members of society as well as communication capabilities.<sup>[3-5]</sup>

Attempts have also been made to identify the human resource characteristics sought by companies using text data. Iwawaki used the text from the message section by the hiring representative among the company information for students.<sup>[6,7]</sup> She coded the vocabulary used in the text and categorized the combinations of vocabulary items using the Quantification III class method to identify the human resource characteristics. Iwawaki identified the vocabulary used in an interview with

a company representative responsible for hiring to deduce the specific abilities that were most highly regarded.<sup>[8]</sup>

Text mining is a comparatively new method of analyzing text data. It is used to extract trends and characteristics from descriptive content by computer-based natural language processing. Ishijima *et al.* analyzed corporate messages aimed at students using text mining.<sup>[9,10]</sup>

Although a large amount of research has been conducted on the resources sought by companies, several issues remain unexplored. First, in Japan, large corporations mainly hire new graduates in bulk. Thus, most existing research only concerns new graduates. Second, the survey methods mainly involve asking a company representative to identify the capabilities that are considered important in candidate selection using a multiple-choice format. However, the conditions that are not anticipated at the time of selection will be missed. Third, majority of the previous research pertains to the abilities considered important by companies in candidate selection rather than investigating whether the candidate was actually hired. Using these research results alone, it is difficult to identify what the companies actually evaluate when making the decision whether to hire a candidate. Even by text mining the data is the message to students during job placements activities, it was unable to deduce the factors that were considered important while making individual selections.

This study attempts to clarify the factors which determined whether an applicant was hired, by text mining on data from individual job referrals for mid-career recruits conducted by Hello Work.

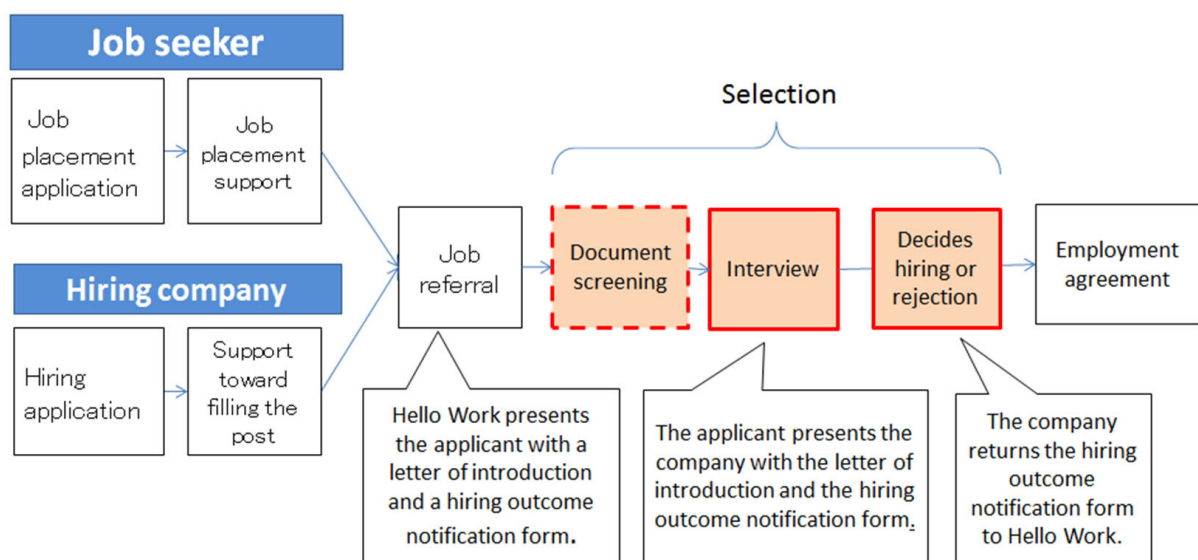


Figure 1. Flow of job referrals in Hello Work

### 3. JOB REFERRALS' DATA FOR ANALYSIS

#### 3.1 The flow of job referrals in Hello Work

Figure 1 shows the flow of job referrals and hire/not hire outcomes in Hello Work. Hello Work receives a job placement

application from the job seeker and provides job placement support, including employment consultations. It also receives hiring applications from companies and in addition to publishing this information, provides support toward filling the post.

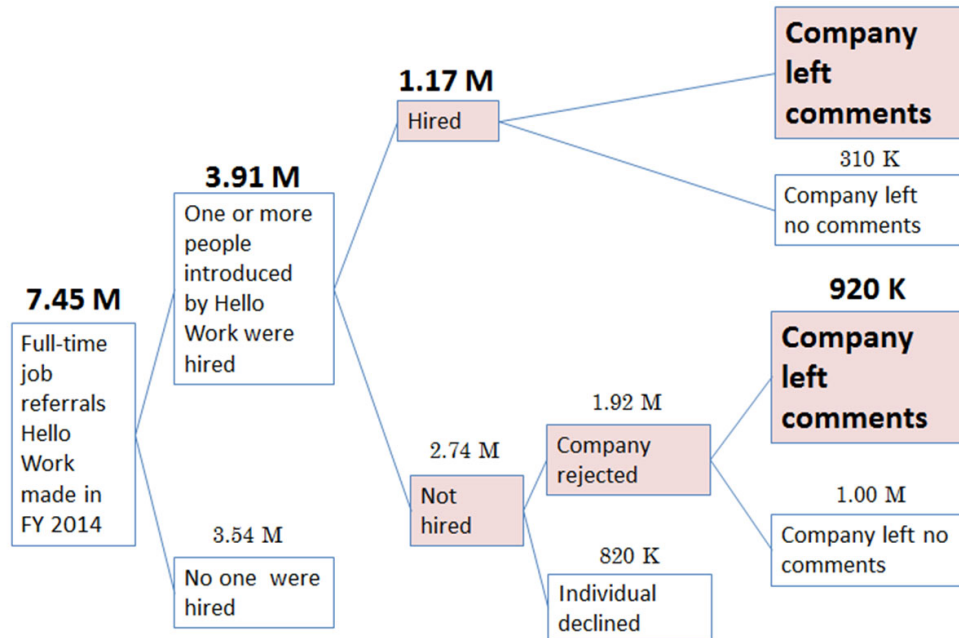


Figure 2. Number of job referrals in Hello Work in FY 2014

Hello Work then matches the job seeker with suitable openings and introduces the job seeker to the hiring company. When it introduces jobs to them, Hello Work provides a letter of introduction and a hiring outcome notification form to the job seeker. The letter introduces the job seeker in question. The hiring company returns the hiring outcome notification form to Hello Work after the outcome is decided.

If the candidate is hired, the company is asked to supply information on the type of occupation, the date of employment commencement, and the salary. If the candidate is not hired, the company is asked to report the reasons for the same. A section in the notification form also allows the company to freely add comments.

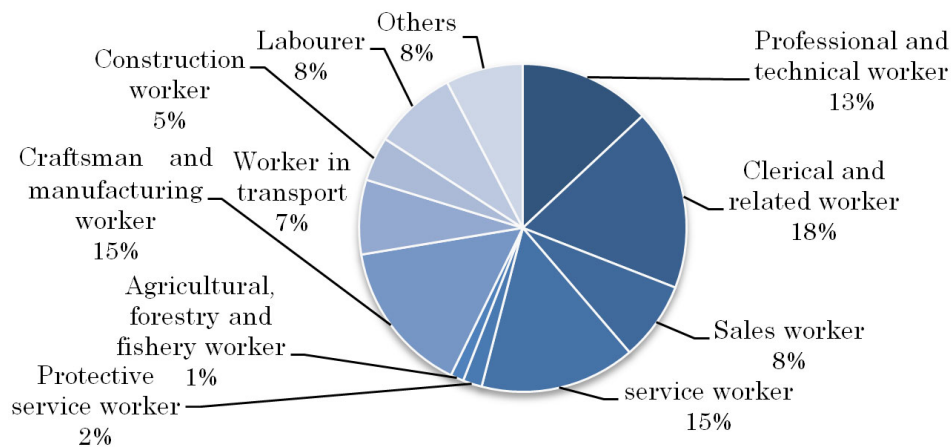
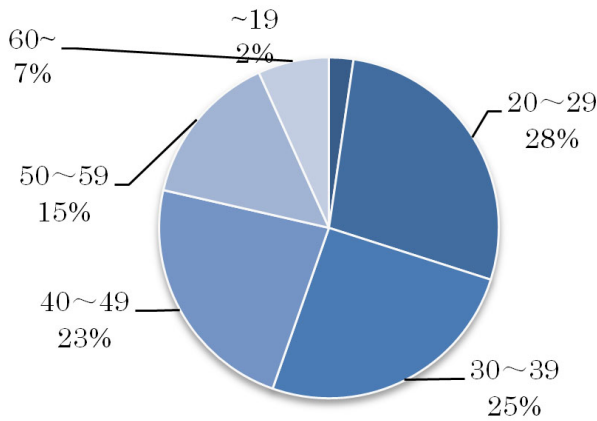


Figure 3. Occupation group ratio of hired data



**Figure 4.** Age group ratio of hired data

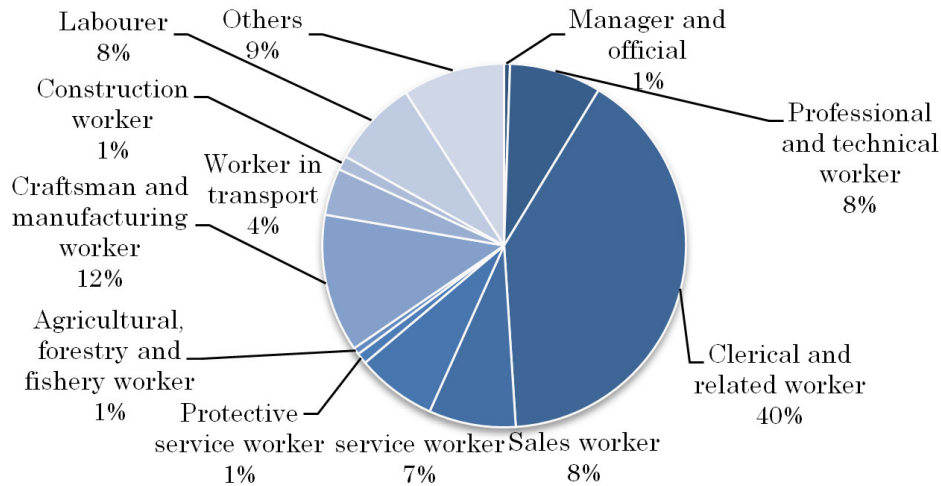
The applicant presents the hiring company the letter of introduction and hiring outcome notification form. The hiring company makes its selection and returns the outcome notification form along with the appropriate details to Hello Work. If the information is not returned, Hello Work follows up

with the hiring company.

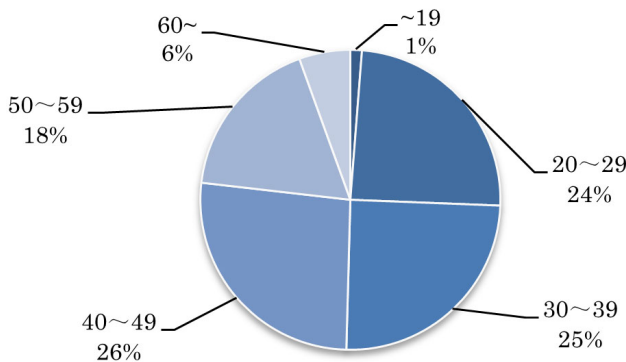
Hello Work has reported the hiring outcomes and the reasons for them, which are used in this study.

**3.2 Analysis using summary values**

Figure 2 shows number of job referrals in Hello Work in FY 2014. Hello Work made approximately 7.45 M full-time job referrals. Of these, approximately 3.91 M job referrals were related to job openings for which one or more people introduced by Hello Work were hired. In these cases, approximately 1.17 M candidates were hired and approximately 2.74 M were not hired. This study analyzed approximately 1.78 M cases for which information on the reasons for hiring or not hiring was entered in the hiring outcome form. Candidates were hired in approximately 860 K of these cases, whereas candidates were not hired in approximately 920 K cases.<sup>[11]</sup> Figures 3-6 show the type of openings and seeker’s attribute.



**Figure 5.** Occupation group ratio of not hired data



**Figure 6.** Age group ratio of not hired data

**4. ANALYSIS USING TEXT MINING OF THE HIRING COMMENTS**

**4.1 Contents and issues of the hiring comments**

Any reasons for hiring or not hiring written in the free comments section by the hiring company are entered into the Hiring Outcome Comments area of the Hello Work database alongside the reasons given over phone calls between the hiring company and Hello Work. In the case of applicants who are hired, details of the employment commencement date and salary are also recorded in this section. Table 1 gives examples of entries in the Hiring Outcome Comments.

**Table 1.** Examples of hiring outcome comments

Group	Comments
Hire	I felt their desire.
	Confirm hiring decision by telephone. Request FAX. April 28 <sup>th</sup> hiring outcome notification.
	April 1 <sup>st</sup> 2014 - hired. Considered to be closest to resource we want to hire. Confirm hiring notification.
	Long experience of sales. Has good philosophy.
	Hiring date - from March 16 <sup>th</sup> 2014 *March 4 <sup>th</sup> send trial plan by post; March 31 <sup>st</sup> will be regular employee from start, so will not use trial.
	Hiring date - April 1 <sup>st</sup> . Does not have experience but is motivated and thinks in a careful way. Decided to hire based on his foundational education (after document selection, decided during interview test).
	Feb 12 <sup>th</sup> - confirmed using hiring outcome notification.
Not hire	Document selection failed.
	At the hospital they worked before, they carried out mainly dispensing work (pharmaceutical assistant) so felt they had limited knowledge of sustained work. Many times, the answers that came back did not follow purport of question. Felt they were lacking in positivity.
	Hired somebody else.
	Comeback was unclear. Confirmed through hiring outcome notification on Jan 31 <sup>st</sup> .
	Because interviewed and hired applicant whose resume arrived first.
	Judged that they did not really match the job.
	Applicant had changed jobs many times and never stayed in same job for long.

As observed, the hiring outcome comments include specialized terms, paperwork-related entries, and orthographic variants. A number of issues must be resolved before attempting to extract the factors that determine hiring or rejection from the entries in the hiring outcome comments.

The first issue is the specialized terms that are included. Some of these terms such as “Tokkaikin (abbreviated name of one of subsidies)” are unfamiliar and seldom used. When performing morphological analysis, these terms may not be processed as single vocabulary items. Therefore, building a dictionary of these unknown terms in advance is necessary.

The second issue is the cases that do not reach the selection stage and are included in the comments. These are cases in which, although the processing related to the introduction was conducted by Hello Work, the job seeker did not apply for the vacancy. If such cases are included, the accuracy of the analysis will be compromised.

The third issue is that the comments are unedited words of the respondents and may differ depending on the hiring company. Terms with the same or similar meanings may be recorded under different notations. If they are analyzed as different terms, the accuracy of the analysis will be compromised.

The fourth issue is that many of the entries reference information other than the hiring decision, which includes hiring results such as “Hiring outcome notification FAX”, entries regarding vacancy such as “continuing the search”, and entries related to the internal processes of Hello Work such as “entered data on benefit department”.

In the next section, we describe our method for resolving these issues and extracting the factors used by the hiring company when deciding to hire/not hire from the comments.

#### 4.2 Analysis using text mining

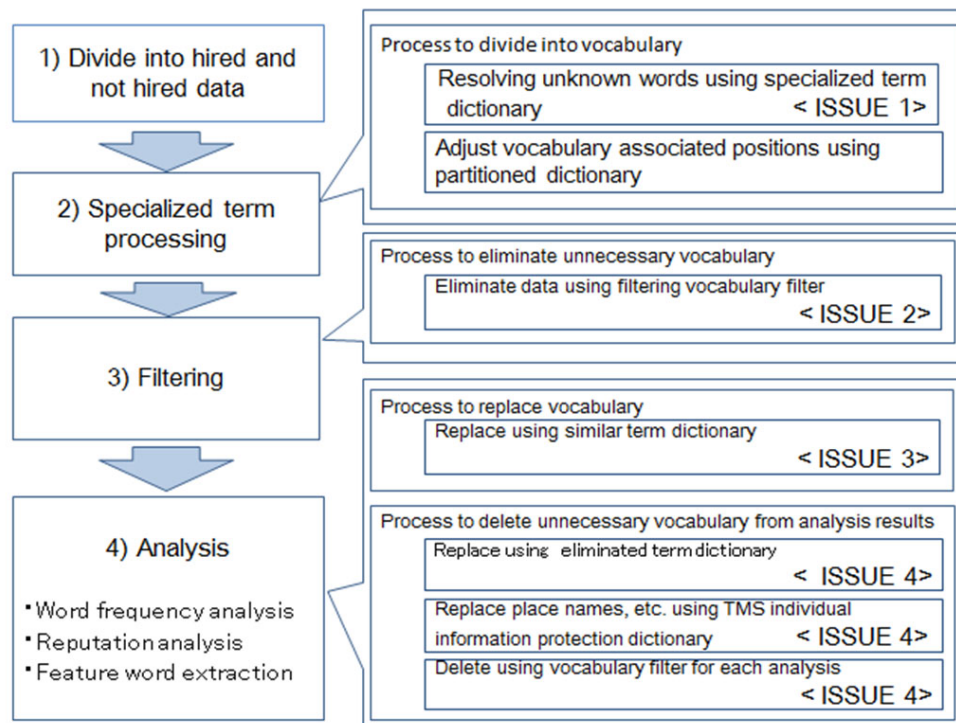
We extracted the hiring or rejection factors using text mining. As the factors of hiring and rejecting decision are not same, and the words to be eliminated which is irrelevant to hiring and not hiring are not same, we divided data into two groups.

In our analysis, we performed natural language processing and eliminated unnecessary data. In our natural language processing, we used Text Mining Studio *ver.* 5.0 for our morphological analyses. Text Mining Studio is a general-purpose text mining tool good at the Japanese analysis that NTT Data Mathematics system Co., Ltd. developed. It is able to complicated analysis by setting and adjusting Text Mining Studio. We categorize comments into morphemes and attached part-of-speech information for each morpheme.

Figure 7 shows analysis flow using text mining. First, we performed a morphological analysis based on the text in the Hiring Outcome Comments. Concretely, we 1) divided hired and not hired data, 2) carried out Specialized term processing, 3) performed filtering, and 4) carried out word frequency analysis, reputation analysis, and feature word extraction.

Chujo and Uchiyama compared the use of eight types of statistical tests (frequency, Dice coefficient, log likelihood ratio, cosine, Yates correction chi-square value, chi-square value, complimentary similarity measure, and self-mutual information quantity) for the detection of feature words and

found that the Dice coefficient is the best individual indicator as it is both effective and valid.<sup>[12]</sup>



**Figure 7.** Analysis flow using text mining

Dice coefficient is the method to quantify the collocation between words and attribute using the number of the appearances and the number of the co-occurrence (see Table 2). Dice coefficient defined by the following formula:

$$Dice\ coefficient = \frac{2a}{(a + b) + (a + c)}$$

**Table 2.** Attribute and word appearance

	Attribute $r_j$	Attribute except $r_j$
Word $w_i$	a	b
Word except $w_i$	c	d

Note. a: Frequency of word  $w_i$  appearing in lines to which attribute  $r_j$  belongs;  
 b: Frequency of word  $w_i$  appearing in lines to which attribute  $r_j$  does not belong;  
 c: Frequency of words other than  $w_i$  appearing in lines to which attribute  $r_j$  belongs;  
 d: Frequency of words other than  $w_i$  appearing in lines to which attribute  $r_j$  does not belong.

Mathematical Systems Inc. <sup>[13]</sup>

For Specialized term processing, a dictionary of specialized terms was created and applied to resolve unknown words (ISSUE 1) in 4.1. We dealt with technical terms related to employment and Hellowork’s abbreviations by registering them in a dictionary (see Table 3). After resolving the unknown words, we created and applied a partitioned dictionary to adjust the positional associations of the vocabulary.

At the filtering stage, vocabulary separation processing was used to exclude irrelevant data, including entries for which job selection had not been completed (such as “Cancel”) (ISSUE 2). After specifying the meaning of each technical term, based on words such as “no show at interview”, we filter out and eliminate data not used in the selection outcome. Table 4 presents examples of words used in the filtering process. After processing, 796,636 hired cases and 893,893 not hired cases remained in the database.

**Table 3.** Example of a technical term dictionary

Notation	Meaning
Toraiaru Koyo (Hiring on a trial basis)	The name of a policy
Kyufu (Benefit)	The name of a department within Hellowork
No show at interview	Used for filtering (no selection took place)

**Table 4.** Filtering list (Examples)

Notation	Meaning
Provisional input	Hiring outcome has not been decided
No show at interview	No selection took place
Cancellation	No selection took place

For analysis, a database of similar vocabulary was created to identify words with the same meaning but different notations (*e.g.*, personal computer/PC) (ISSUE 3).

Finally, we deleted the irrelevant vocabulary using the dictionary of eliminated words, Text Mining Studio Version 5.0 individual information protection database, and a number of

vocabulary filter functions (ISSUE 4).

Next, we preprocessed words to analyze the frequency of each word and the frequency of its associations.

Specifically, we grouped synonyms and synonymous words having different notations, such as “personal computers” and “PCs” or “qualification” and “license”, into a synonym dictionary. Next, we removed from the outcome frequently used words related to “hiring outcome” and “hiring condition”, words used often in the selection process. We also grouped words related to administrative process within Hellowork as they appear in many sentences regardless of the hiring rationale. Table 5 shows the dictionary of eliminated words.

**Table 5.** Dictionary of eliminated words

Word	Meaning
Recruit(ment)	A word often used in the selection process
HW	The abbreviation for Hellowork. A word often used in the selection process
Pass	A word linked to or synonymous with the hiring outcome
Completed copying	A word related to Hellowork paperwork
FAX	A word related to Hellowork paperwork

## 5. ANALYSIS RESULTS USING TEXT MINING

### 5.1 Analysis of text mining of hired cases

When looking at the word frequency analysis results for hired cases, “experience” was the most frequent item, but words related to attitude and character such as “motivation”, “serious”, “desire”, “character”, “bright”, and “sincere” were also highly ranked.

Next, a reputation analysis was performed using the Text Mining Studio Version 5.0 function. With the reputation analysis, positive points are assigned to words that have favorable evaluation such as “good”, “sufficient”, etc. and negative points to words that have unfavorable evaluation such as “lack”, “insufficient”, etc.

**Table 6.** Separating groups of words with a high frequency of use

Symbol	Word	Symbol	Word
【E】	Experience	【S】	Serious, Sincere
【M】	Motivation, Aspiration, Positive, Eagerness	【PH】	Physical ability, Physical condition, Healthy
【AG】	Age	【K】	Knowledge, Skill
【PE】	Personality, Character	【R】	Readiness to work
【AT】	Attitude, Posture	【CO】	Commuting distance, Commuting time
【Q】	Qualification	【CH】	Job change
【V】	Vigorous, Cheerful	【I】	Impression, Atmosphere

To identify characteristics other than experience that were highly regarded, we performed a word frequency analysis. In the analysis, we used the Weblio similar terms dictionary to separate words that appeared with a high frequency as shown in Table 6. These words were tagged with a symbol in the analysis results diagram and table.

Words with high positive values included “impression”, “posture”, “experience”, “character”, “aspiration”, and “motiva-

tion”. While “impression” and “posture” had few negative evaluations, “experience” had both positive and negative evaluations like “Hired. He has experience in sales”, “The experience is poor, but is motivated” . (See Table 7 and Table 8. For further information see Asano<sup>[14]</sup>)

### 5.2 Analysis of text mining of not hired cases

Word frequency analysis results for the not hired cases revealed many terms related to the existence of better can-

didates such as “applicant”, “many”, “others”, “qualified person”, and “comparison”. Other frequent words such as “comprehensive” did not provide detailed reasons. Although it is more difficult to deduce the factors for not hiring than those for hiring, experience, age, and skill were key factors.

Next, when looking at the reputation analysis results for not hired cases, words with low evaluation included “job change”, “experience”, “age”, “commuting distance”, and “physical ability”. While some terms such as “job change” and “physical side” had a purely negative evaluation, others had both positive and negative evaluations, including “experience”, “age”, and “knowledge”, for example, “She has 2 years’ experience, but has no qualifications”, “Because of poor experience, we can’t hire him”, *etc.* ( See Table 9 and

Table 10. For further information see Asano<sup>[15]</sup> ).

**Table 7.** Word frequency analysis results for “hired” data

Ranking	Symbol	Word	Frequency
1	<b>【E】</b>	Experience	11,741
2		Operations	9,735
3	<b>【M】</b>	Motivation	6,034
4		Good	5,757
5	<b>【S】</b>	Serious	4,495
6	<b>【M】</b>	Aspiration	4,465
7	<b>【PE】</b>	Personality	3,760
8	<b>【Q】</b>	Qualification	2,864
9	<b>【V】</b>	Vigorous	2,534
10	<b>【R】</b>	Readiness to work	2,430

**Table 8.** Reputation analysis results for “hired” data

Ranking	Symbol	Word	Positive	Negative	Positive-Negative
1	<b>【I】</b>	Impression	1,125	11	1,114
2	<b>【AT】</b>	Posture	914	2	912
3	<b>【E】</b>	Experience	1,137	285	852
4	<b>【PE】</b>	Personality	729	5	724
5		Operations	604	125	479
6		Courteous	438	0	438
7	<b>【S】</b>	Serious	346	12	334
8	<b>【M】</b>	Aspiration	315	8	307
9	<b>【AT】</b>	Attitude	267	9	258
10		Smile	252	0	252

**Table 9.** Word frequency analysis results for “not hired” data

Ranking	Symbol	Word	Frequency
1		Applicant	36,329
2		Many	28,355
3		Operations	22,662
4	<b>【E】</b>	Experience	19,546
5		Other person	12,651
6	<b>【AG】</b>	Age	12,208
7		Good	10,390
8		Well-qualified person	9,551
9		Difficult	8,489
10		Different	8,075

**5.3 Observations on the results for hired and not hired data**

From these results, it is observed that there are 1) words that appeared with high frequency in both the hired and not hired data, such as “experience”, “knowledge”, and “qualifications”, 2) words that appeared frequently in the hired data but were less common in the not hired data, such as

“motivation”, “serious”, “personality”, and “posture”, and 3) words that appeared with high frequency in the not hired data but were less common in the hired data, such as “applicant” and “job change”.

Furthermore, it was clear from the analysis that both the hired and not hired data contained a mix of positive and negative evaluation words.

We focused particularly on the word “experience”, which appeared with high frequency. We found that when positive evaluation words were included in the hired data, the evaluation of experience was negative (270 cases) and when negative evaluation words were included in the not hired data, the evaluation of “experience” was positive (483 cases). We used these to search for factors other than experience, which determined hiring or rejection.

**5.4 Analysis of hired data for which the evaluation of experience was negative**

In the advanced analysis, to check whether the attribute of data does not have deflection, we compared the attributes



of hired data for which the evaluation of experience was negative (270 cases) with the total hired dataset (862,160 cases). The results showed that the hired data in which the evaluation of experience was negative had the following characteristics: 1) many cases were in the 20-29 years age group, and, 2) many cases involved companies with nine or fewer

employees.

Next, we analyzed the application condition of the job offer of the data which evaluation of experience was negative, but decided to be hired by other factors. Previous to analyze, it is useful to know the application condition required the experience or not.

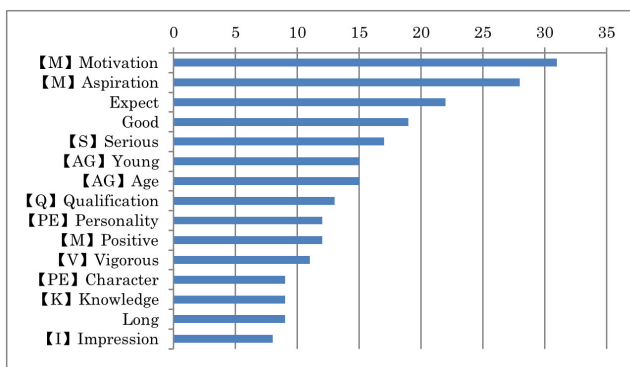
**Table 10.** Reputation analysis results for “not hired” data

Ranking	Symbol	Word	Positive	Negative	Positive-Negative
1	【CH】	Job change	7	2,257	-2,250
2	【E】	Experience	1,124	2,547	-1,423
3	【AG】	Age	180	1,519	-1,339
4	【CO】	Commuting distance	43	385	-342
5	【PH】	Physical side	10	312	-302
6	【R】	Readiness to work	81	307	-226
7		Communication skills	61	282	-221
8		Reason for applying	31	243	-212
9	【PH】	Health	4	193	-189
10	【AT】	Attitude	52	229	-177

When analyzing cases in the hired data for which the evaluation of experience was negative, openings that required experience were 40.4% and those not requiring experience were 59.6%. In cases where experience was not required, it is not surprising that some applicants were hired despite being negatively evaluated in terms of experience. However, in 40% of the cases, candidates were hired even though experience was required and they were negatively evaluated in terms of experience.

Next, we performed feature word extraction.

In hired data for which the evaluation of “experience” was negative, the results of feature word extraction based on cases where experience was and was not required are shown in Table 11. It is observed that for both types of openings, aspiration, motivation, seriousness, and age were considered to be important. On the other hand, qualification was considered to be important in the job offer of the experience was not required.



**Figure 8.** Word frequency analysis results from hired data for which the evaluation of “experience” was negative

The word frequency analysis results for hired data in which experience was evaluated negatively are shown in Figure 8. It can be concluded that “seriousness” and “youth” were highly regarded in addition to “motivation” and “aspiration”.

**Table 11.** Feature word extraction results by necessity/non-necessity of experience in hired data for which the evaluation of “experience” was negative

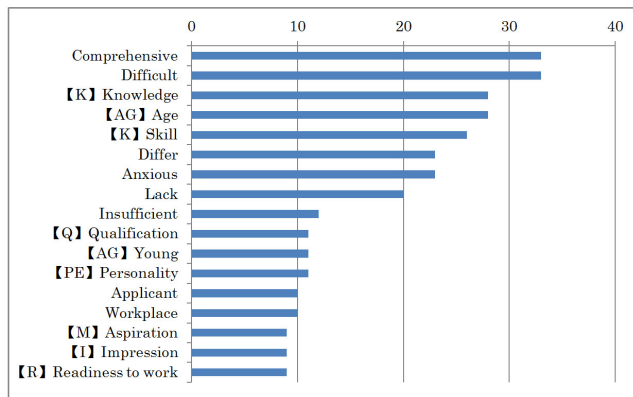
Ranking	Experience is not required	Experience is required
1	【M】 Motivation	【M】 Aspiration
2	【M】 Aspiration	Expect
3	Good	【AG】 Young
4	Expect	【S】 Serious
5	【Q】 Qualification	【M】 Motivation
6	【S】 Serious	【AG】 Age
7	【PE】 Personality	Good
8	【AG】 Age	【M】 Positive
9	【V】 Vigorous	【V】 Vigorous
10	【M】 Positive	【PE】 Personality

**5.5 Analysis of not hired data for which the evaluation of “experience” was positive**

Prior to the analysis, to check whether the attribute of data does not have deflection, we compared the attributes of not

hired data for which the evaluation of experience was positive (483 cases) with the total not hired dataset (919,685 cases). The results showed that the not hired data in which the evaluation of experience was positive had the following characteristics: 1) many cases were in the 40 or over age group, 2) many cases involved specialist technical work or office work, whereas few were in production process work, transportation equipment delivery, or cleaning and packaging, and 3) many cases involved companies with nine employees or fewer.

Further, regarding to the job openings where the candidate was not hired and the evaluation of experience was positive, 53.4% openings required experience, whereas 46.4% did not. This suggests that higher experience is required in such cases rather than where the evaluation of experience was negative yet the candidate was hired.



**Figure 9.** Word frequency analysis results in not hired data for which the evaluation of “experience” was positive

**Table 12.** Feature word extraction results by necessity/non-necessity of experience in hired data for which the evaluation of “experience” was negative

Ranking	Experience is not required	Experience is required
1	Comprehensive	【K】 Knowledge
2	【AG】 Age	【K】 Skill
3	Difficult	Differ
4	Anxious	Difficult
5	Lack	Lack
6	【K】 Skill	Anxious
7	【PE】 Personality	【AG】 Age
8	Differ	Insufficient
9	【K】 Knowledge	Workplace
10	Applicant	【Q】 Qualification

A word frequency analysis was conducted to identify the factors leading to a candidate not being hired despite a positive

evaluation of experience. The results are shown in Figure 9. The results show that this includes cases where despite the fact that experience, knowledge or skill was sufficient, there were problems in terms of age or the experience of the candidate in a different field.

Next, when analyzing word frequency by whether experience was or was not required, the results are shown in Table 12. This shows that for openings in which experience was not required, age and personality were highly valued, whereas for openings in which experience was required, knowledge and skill were highly valued.

## 6. CONCLUSION

This study aimed to investigate the factors that determine hiring or rejection, by applying text mining to the comments made by hiring companies.

The results showed that in the hired cases, vocabulary related to character such as “serious” and “personality” appeared frequently in addition to “experience” and “aspiration”. The not hired data contained many vocabulary items related to rival candidates, and it was difficult to grasp these factors. In addition to “job change” and “experience”, we extracted terms related to age and commuting. Terms such as experience appeared in cases of both hired and not hired. This shows that positive evaluations such as aspiration or personality, and negative evaluations such as job change were factors that could be easily linked.

Further, as in the case of experience, a single evaluation is not necessarily linked to the hiring outcome as certain factors allow the situation to be retrieved even when a negative evaluation is made. Moreover, when focusing on experience and investigating data where the candidate was hired despite lacking experience, 40% or more of the openings were ones that required experience. This shows that hiring companies value other factors highly in addition to experience.

Additionally, from data in which the candidate was not hired despite having experience, the main reasons seemed to be that their experience was in a different field.

The study has added formal knowledge about the previously tacit factors that determine hiring decisions in Japan. The approach taken in this study is not only academically significant in terms of both method and analysis, but can also contribute in improving the efficiency with which positions are matched to people as well as strengthening the external labor market.

Moving forward, we will build on this analysis and clarify the differences based on occupation and age.

**REFERENCES**

- [1] General Status of Job Referrals (Monthly). Internet. Ministry of Health, Labour and Welfare [cited 2016 April 30]. Available from: <http://www.e-stat.go.jp/SG1/estat/List.do?lid=000001150335>
- [2] Hamaguchi K. Youth and labor. Chuoukouronshinsha; 2013.
- [3] Ministry of Economy Trade and Industry. Emergency Survey results of fundamental professional abilities; 2006.
- [4] Japan Business Federation. Summary of a Questionnaire Regarding Recruitment of New Graduates (for April 2014 hirings); 2014.
- [5] Japan Association of Corporate Executives. Results of "Questionnaire survey regarding corporate hiring and education" (2014 survey); 2014.
- [6] Iwawaki C. Research on the Portrait of 'the preferred individual': Using a dual comparison of remarks made by a renowned corporation. *Research on Educational Sociology*. 2004; 74: 309-27. <http://dx.doi.org/10.11151/eds1951.74.309>
- [7] Iwawakii C. Transformation in the image of desired human resources in new university graduate hiring since the advanced growth period. *Kyoto University Graduate School Education & Research Dept. Bulletin*. 2006; 52: 79-92.
- [8] Iwawaki C. Structure and Changes in 'Capabilities' Sought for in a New College Graduate. *Works Review*. 2006; 1.
- [9] Ishijima S, Udagawa Y. Analysis of Help Wanted Ads with the Use of Data Mining: with the hope of decreasing mismatches between corporations and students. *Shingaku Gihou*. IEICE Technical Report SWIM. 2012; 17: 13-8.
- [10] Ishijima S, Udagawa Y. Examples of Help Wanted Ad Analyses Using Co-occurrence of Words, 75th National Convention of the Information Processing Society of Japan. 2013; 1: 513-4.
- [11] Main initiative and results of Hello Work. Internet. Ministry of Health, Labour and Welfare; 2014. [cited 2016 May 12]. Available from: <http://www.mhlw.go.jp/file/06-Seisakujouhou-11600000-Shokugyouanteikyoku/0000067862.pdf>
- [12] Chujo K, Uchiyama M. Using Statistical Measures to Extract Specialized Vocabulary from a Corpus. *Kantokoshinetsu Association of Teacher of English (KATE)*. 2004: 99-108.
- [13] Mathematical Systems Inc. *Text Mining Studio Version 5.0 Technical Materials*; 2014.
- [14] Asano H. Study on Extraction of Hiring Determinant from the Job Placement Offices Data using Text Mining. *ICBIS2015 (International Conference on Business, Information, and Service Science)*; 2015. PMID:26502663.
- [15] Asano H. Study on hiring decision: analyzing rejected applicants by mining individual job placement data of Public Employment Service Offices. *KES 2015 (19th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems)*. 2015: 1156-63. PMID:26502663.