



SENTIMENT ANALYSIS METHOD FOR JAPANESE QUESTIONNAIRES USING JAPANESE–ENGLISH TRANSLATION

Kyoko Kabasawa ^a, Kazuhiko Tsuda ^b

^a University of Tsukuba, Tokyo Japan
kabasawa.kyoko.ww@alumni.tsukuba.ac.jp

^b University of Tsukuba, Tokyo Japan

Corresponding Email: tsuda.kazuhiko.gn@u.tsukuba.ac.jp

Abstract

In sentiment analysis using a dictionary, words that form a sentence are extracted by a morphological analysis for each part of speech and the sentence is judged to be positive or negative based on the bipolarity values of the words that appear in it. Therefore, even if a sentence is translated from Japanese to English, the polarity of the entire sentence remains unchanged. The purpose of this study was to analyze the contents of questionnaires to obtain a questionnaire with negative information. If the business owner can determine particularly negative points from the reviews to identify the scope for improvement, it may be possible to streamline the work done by individuals who consider each review. We analyzed whether the content of questionnaire data obtained from customers who used restaurants was positive or negative. We compared the method of using the Japanese polarity dictionary with the method of translating the questionnaire into English and using the English polarity dictionary, and discussed which method was able to determine polarity more predictably. In terms of negative precision, an indicator of whether the negative content was obtained as predicted, English analysis outperformed. Also, we found that the polarity of a sentence depends on the number of words in the polarity dictionary. Furthermore, if it is difficult to determine the polarity of an entire sentence, it is better to determine the polarity of each sentence using Aspect Based Sentiment Analysis.

Keywords: insert sentiment analysis, polarity analysis, classification evaluation, natural language processing, text mining

1. Introduction

In this study, we analyze questionnaires in the restaurant industry and investigate how they can be beneficial to business owners. Specifically, the review portion of the questionnaire by the contributors and the score assigned by them to the target restaurant were compared, and it was determined whether there was any discrepancy between the review and the score. If the business owner can determine particularly negative points from the reviews to identify the scope for improvement, it may be possible to streamline the work done by individuals who consider each review and make a business decision accordingly. Therefore, the purpose of this study was to analyze the contents of questionnaires to obtain a questionnaire with negative information.

2. Research problem

In natural language processing, Japanese is considered to be a more difficult language than English for analysis, owing to two reasons. The first reason is that Japanese sentences do not have spaces between words, as in English sentences. In Japanese, because each part of speech is concatenated to form a sentence, it a machine should learn the regularity of the language to understand the meaning of a sentence by correctly identifying and separating the clauses. The second reason is the difficulty in understanding the conclusions, which is unique to Japanese. In English, the subject is followed by the predicate and then the object or the modifier; thus, the intent can be understood at the beginning of the sentence. In contrast, in Japanese, the object or the modifier is embedded between the subject and the predicate; thus, it technically requires more steps to mechanistically determine the intent of the sentence compared to the steps required for English. In general, when the narrow-sense semantic analysis of Japanese sentences using the sentiment approach is compared with the broad-sense semantic analysis of Japanese sentences that are machine-translated from English, the accuracy is low. However, in sentiment analysis using a dictionary, words that form a sentence are extracted by a morphological analysis for each part of speech (noun, verb, adjective, and adverb), and the sentence is judged to be positive or negative based on the bipolarity values of the words that appear in it. Therefore, even if a sentence is translated from Japanese to English, the polarity of the entire sentence remains unchanged. Accordingly, in this study, the difference between the analyses of questionnaire results obtained in Japanese and their English translations was examined for ease of judgment for a machine in an attempt to obtain positive and negative information from the questionnaire results.

3. Review of the relevant literature

Wilson, Wiebe and Hoffmann (2005) proposed a sentiment analysis method to determine whether a phrase is polar or neutral and, if it is polar, to classify the polarity according to the polarity of both words and contexts. Shinzato and Oyamada (2018) considered a sentential perspective and conducted a polarity assessment to classify each sentence in the free-text reviews of online shopping stores so that users could effectively access desired information. Nanya and Maeda (2019) proposed a method of recommending restaurants to foreigners according to their preferences by using word-of-mouth sites. They considered the structure and sentiment involved in the review and included contextual data, such as the place of origin, in their analysis. Thus, the results of sentiment analysis have been applied in various fields.

4. Method

Data were obtained from questionnaires that were completed and submitted by customers after they visited a restaurant and received service as mystery shoppers. The data comprised 2607 responses to the following two questionnaire items: "Please provide a general comment on the service and food of the restaurant you used this time" and "Please provide the overall score for the food and customer service."

First, comments of 2607 target data were translated into English. In addition to Google Translate, Gengo.com (2021) was used to translate them into English. Then, polarity analysis based on positive and negative reviews was performed. The target Japanese and English data were subjected to polarity analysis using Positive-Negative Table (hereinafter referred to as PNTTable) (Takamura, Inui and Okumura, 2005) (Fig. 1), which is based on the Iwanami Japanese Dictionary (Iwanami Shoten) for Japanese and WordNet-1 7.1 for English. In addition, the SentimentAnalysis (hereinafter referred to as SA) (2021) package in R was used. The Harvard IV-4 categories of polar dictionaries (2021) were used. As shown in Equation (1), the polarity value of a sentence using PNTTable is a sum of the polarity value of each word in each polarity dictionary (with polarity value) multiplied by its frequency of occurrence.

$$X_{pn} = \sum_{i=0}^n (ki) \quad (1)$$

X_{pn}: Polarity of the entire sentence
 k: Polarity value of the word
 i: Frequency of occurrence of the word

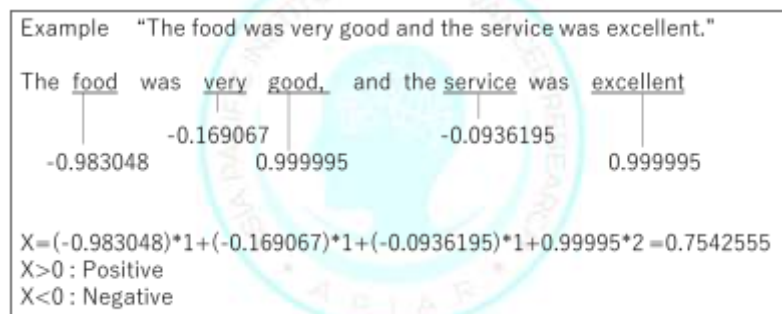


Figure 1: Determining the polarity of a sentence (Actual example sentences are written in Japanese.)

Meanwhile, the polarity value of a sentence using SA was obtained as positive (1), neutral (0), or negative (-1). If words in a sentence exist in the dictionary, their polarity values are added and divided by the number of words that occur in the sentence, as shown in Equation (2).

$$X_{sa} = \frac{p+(-1)*n}{p+n} \quad (2)$$

X_{sa}: Polarity value of the entire sentence
 n: Number of occurrences of positive-polarity words
 p: Number of occurrences of negative-polarity words

For the evaluation of a restaurant, the score given in response to the questionnaire item "Please provide the overall score for the food and customer service" was analyzed. The score was given in 10-point increments on a 100-point scale. Therefore, the score range of 0 to 50 was labeled as

negative, 60 to 70 as neutral, and 80 to 100 as positive, and these labels were used as the baseline values to determine the polarity of Japanese sentences and their English counterparts.

Meanwhile, manual annotation was performed, and the results were analyzed. The annotations were divided into six categories: positive, negative, neutral, request, negative + positive, and difficult to determine. Positive annotation was given when the meaning of the sentence was positive and there were no negative words, and, conversely, negative annotation was given when the meaning of the sentence was negative and there were negative words. Neutral annotation was given when the sentence only conveyed information without emotion, such as "This store opens at 5:00 p.m." Request annotation was given when the request was explicit, such as "I would be happy if you could do that," "I want that," or "I think it would be good if you could do that." Negative + positive annotation was given when the sentence included at least one positive and one negative emotion, and difficult-to-distinguish annotation was made when the sentence was extremely long for the reader to determine whether it represented a negative or positive emotion. Figures 2–4 illustrate the polarity values of Japanese sentences and English counterparts for 1 to 50 reviews.



Figure 2: Polarity of Japanese sentences (PNTTable)

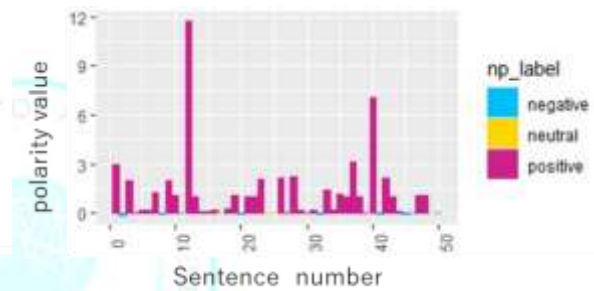


Figure 3: Polarity of English counterparts (PNTTable)

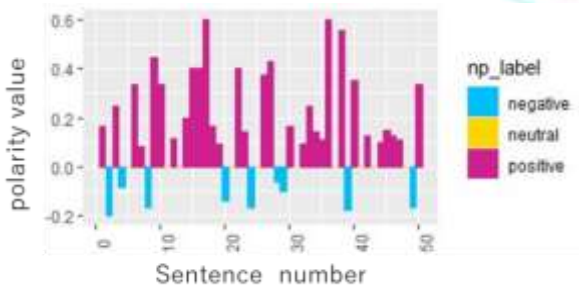


Figure 4: Polarity of English counterparts (SA)

5. Data analysis

Score-based data showed that the reviews were positive in 1622 cases, negative in 232, and neutral in 753. In the analyses, neutral responses were excluded.

Tables 1–3 show the number of judgements by each analysis based on the score-based results.

Table 1: Score-based judgment of Japanese sentences

Score-based Japanese(PNTable)	
Positive→	Negative→
Judgement:Positive (TruePositive)	Judgement:Positive (FalsePositive)
728	37
Positive→	Negative→
Judgement:Negative (FalseNegative)	Judgement:Negative (TrueNegative)
804	187

Table 2: Score-based judgement of English counterparts

Score-based English(PNTable)	
Positive→	Negative→
Judgement:Positive (TruePositive)	Judgement:Positive (FalsePositive)
1277	186
Positive→	Negative→
Judgement:Negative (FalseNegative)	Judgement:Negative (TrueNegative)
251	39

Table 3: Score-based judgement of English counterparts (SA)

Score-based English(sentimentAnalysis)	
Positive→	Negative→
Judgement:Positive (TruePositive)	Judgement:Positive (FalsePositive)
1154	129
Positive→	Negative→
Judgement:Negative (FalseNegative)	Judgement:Negative (TrueNegative)
196	61

Tables 4–6 show the rates of accuracy, precision, recall, specificity, and F-measure for Japanese and English.

Table 4: Score-based evaluation of classification prediction (Japanese)

Japanese(PNTable)					
Accuracy	0.5211				
precision (positive)	0.9516	recall (positive)	0.4752	F-measure (positive)	0.6339
precision (negative)	0.1887	recall (negative)	0.8348	F-measure (negative)	0.3078

Table 5: Score-based evaluation of classification prediction (English)

English(PNTable)

Accuracy 0.7504					
precision (positive)	0.8729	recall (positive)	0.8357	F-measure (positive)	0.8539
precision (negative)	0.1345	recall (negative)	0.1733	F-measure (negative)	0.1515

Table 6: Score-based evaluation of classification prediction (English SA)

English(sentimentAnalysis)

Accuracy 0.7890					
precision (positive)	0.8995	recall (positive)	0.8548	F-measure (positive)	0.8766
precision (negative)	0.2374	recall (negative)	0.3211	F-measure (negative)	0.273

Annotation-based analysis showed that positive annotation was found in 1273 cases, negative annotation in 415 cases, neutral annotation in 49 cases, request annotation in 108 cases, negative + positive annotation in 745 cases, and difficult-to-distinguish annotation in 17 cases. Of these, 1688 responses with a positive or negative annotation were used for the analysis. Reviews that were found to be neutral in the course of each analysis were excluded.

Tables 7–9 show the number of judgments made by each analysis based on annotation-based results.

Table 7: Annotation-based judgment of Japanese sentences

Annotation-based Japanese(PNTable)			
Positive→		Negative→	
Judgement:Positive (TruePositive)		Judgement:Positive (FalsePositive)	
	716		17
Positive→		Negative→	
Judgement:Negative (FalseNegative)		Judgement:Negative (TrueNegative)	
	470		380

Table 8: Annotation-based judgement of English counterparts

Annotation-based English(PNTable)			
Positive→		Negative→	
Judgement:Positive (TruePositive)		Judgement:Positive (FalsePositive)	
	985		311
Positive→		Negative→	
Judgement:Negative (FalseNegative)		Judgement:Negative (TrueNegative)	
	185		90

Table 9: Annotation-based judgement of English counterparts (SA)

Annotation-based English(sentimentAnalysis)	
Positive→ Judgement:Positive (TruePositive)	Negative→ Judgement:Positive (FalsePositive)
960	196
Positive→ Judgement:Negative (FalseNegative)	Negative→ Judgement:Negative (TrueNegative)
109	124

Tables 10–12 show the rates of accuracy, precision, recall, specificity, and F-measure for Japanese and English.

Table 10: Annotation-based evaluation of classification prediction (Japanese)

Japanese(PNTable)

Accuracy 0.6924					
precision (positive)	0.9768	recall (positive)	0.6037	F-measure (positive)	0.7462
precision (negative)	0.4471	recall (negative)	0.9572	F-measure (negative)	0.6095

Table 11: Annotation-based evaluation of classification prediction (English)

English(PNTable)

Accuracy 0.6843					
precision (positive)	0.76	recall (positive)	0.8419	F-measure (positive)	0.7989
precision (negative)	0.3273	recall (negative)	0.2244	F-measure (negative)	0.2663

Table 12: Annotation-based evaluation of classification prediction (English SA)

English(sentimentAnalysis)

Accuracy 0.7804					
precision (positive)	0.8304	recall (positive)	0.898	F-measure (positive)	0.8629
precision (negative)	0.5322	recall (negative)	0.3875	F-measure (negative)	0.4485

6. Discussion

In the actual analysis of questionnaires, it is not realistic to review the content of a large number of questionnaires. Therefore, it is common to review the low scores given to the request "Provide the overall score for the food and customer service" and collect negative comments to the request "Please provide a general comment on the service and food of the restaurant you used this time." However, this method requires a large amount of time to review a large number of questionnaires. Therefore, it is essential to be able to obtain negative feedback as predicted from the questionnaire without having to review all the questionnaires. Hence, we investigated the effectiveness of precision on negative comments.

First, comparing the score-based and annotated-based approaches, 1190 (approximately 46%) data items out of 2607 data items were deemed to share the same polarity value (positive, neutral, or negative) in both approaches. Thus, score-based analysis of both English and Japanese text revealed a gap between the score and the comment, indicating that correct positive and negative judgments were not made. This was also reflected in the results for accuracy in the analysis of Japanese text.

In the annotation-based analysis, two analyses of English text were compared first. In the analysis using the English PNTTable, the number of occurrences of words that were also included in the dictionary was 6920 from 1688 documents, whereas, in the analysis using SA, the number of occurrences of words that were also included in the dictionary was 22,242 out of 1688 documents, a threefold difference. Because the dictionary used was not field-specific, it was expected that the words contained in the dictionary, especially those used in the English PNTTable, would not match well with the words in the documents. In the Japanese PNTTable, the number of occurrences of the words that were also included in the dictionary was 18,687 out of 1688 documents. Therefore, the annotation-based analysis using SA for English text and the analysis using the Japanese PNTTable for Japanese text were compared.

The F-measure for positive comments was higher in the English analysis, but the F-measure for negative comments was higher in the Japanese analysis. However, this is thought to result from the bias in the difference between precision and recall in negative comments in Japanese, resulting in a larger F-measure. In terms of negative precision, an indicator of whether the negative content was obtained as predicted, English analysis outperformed Japanese analysis. We suggest that method to translating Japanese text into English as to get negative information from them.

Conclusion

We proposed a technique for determining the positivity or negativity of review comments in a questionnaire survey. This technique was evaluated by translating Japanese comments into English and then determining their positivity or negativity. In this analysis, the results annotated with "request" or "positive + negative" were not analyzed. It is believed that perspectival sentiment analysis on these data will make it possible to determine positive or negative perspectives and themes more specifically. These results can be used to improve business strategies further.

Acknowledgment

The data for this study was provided by ROI Inc. We would like to thank them.

Rereferences

- i. Wilson, T., Wiebe, J., and Hoffmann, P., Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis., *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, Vancouver, Canada, October 6-8, 2005.
- ii. Shinzato, K. and Oyamada, Y. 2018. What Do People Write in Reviews for Sellers? -Investigation and Development of an Automatic Classification System-. *Natural Language Processing* 25(1), 57-79.
- iii. Nanya, Y., and Maeda, A., Using word-of-mouth sites to recommend restaurants for foreign visitors to Japan, The 11th Forum on Data Engineering and Information Management, Nagasaki, Japan, March 3-6, 2019.
- iv. Gengo, available at: <https://gengo.com/ja/> (accessed 22 April 2021)
- v. Takamura, H., Inui, T., and Okumura, M., Extracting Semantic Orientations of Words using Spin Model, *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics*, Ann Arbor, USA, June 25-30, 2005
- vi. SentimentAnalysis, available at: <https://cran.r-project.org/web/packages/SentimentAnalysis/SentimentAnalysis.pdf> (accessed 22 April 2021)
- vii. Harvard dictionary, available at: <http://www.wjh.harvard.edu/~inquirer/homecat.htm> (accessed 22 April 2021)

