Keyword Extraction for Local Foods from Restaurant Menus of Roadside Stations

*Yoko Nishihara (presenter), *Hirofumi Noguchi, and **Ryosuke Yamanishi

* College of Information Science and Engineering, Ritsumeikan Univ.

** Faculty of Informatics, Kansai University

nisihara@fc.ritsumei.ac.jp

Short resume of the presenter

- Name: Yoko Nishihara
- Title: Professor (Dr. of Engineering)
- Affiliation: College of Information Science and Engineering, Ritsumeikan University, Japan
- My laboratory's Web site: https://www.nisihara-lab.org/

Topics of research interest of our group

- Human-Computer Interaction: https://tinyurl.com/2mnxhbmw (demo)
- Natural Language Processing: https://tinyurl.com/2yspn5k8 (slides)
- Comic Computing
- Edutainmeint
- Entertainment: https://tinyurl.com/r49366sb (demo)
- Multimedia on Cooking and Eating Activities: https://tinyurl.com/8vend223 (slides)

Research backgroud (1)

Tianjin Xiao Long Bao (China)

- People enjoy not only sightseeing on their vacation but also food tourism.
- They enjoy local foods that are eaten in specific regions.
- Natives for the region may be confused which foods are local.
  - They are accustomed to these foods.

B. C. roll (Canada)

Nanaimo bar (Canada)

which food is local for this region?

Food A

Food B

Food C

https://en.wikipedia.org/wiki/Xiaolongbao
https://en.wikipedia.org/wiki/Nanaimo_bar
https://en.wikipedia.org/wiki/B.C._roll

Research background (2)

- Further studies have been conducted to discover local foods automatically by applying statistical methods to restaurant menus [Tanaka, Yamanishi, and Nishihara 2018] (in Japanese).
  - They applied their method to restaurant menus on a Web site, Yelp.
  - The restaurant menus do not necessarily include local foods.

https://www.kagaminosato.com/

Objective of this paper

- We apply the previous statistical method to **restaurant menus in Japanese roadside stations**.
  - As the roadside stations provide foods produced in those regions, the restaurants may offer menus that include the local foods.

ingredients provided at a roadside staton

a menu offered
at the roadside station

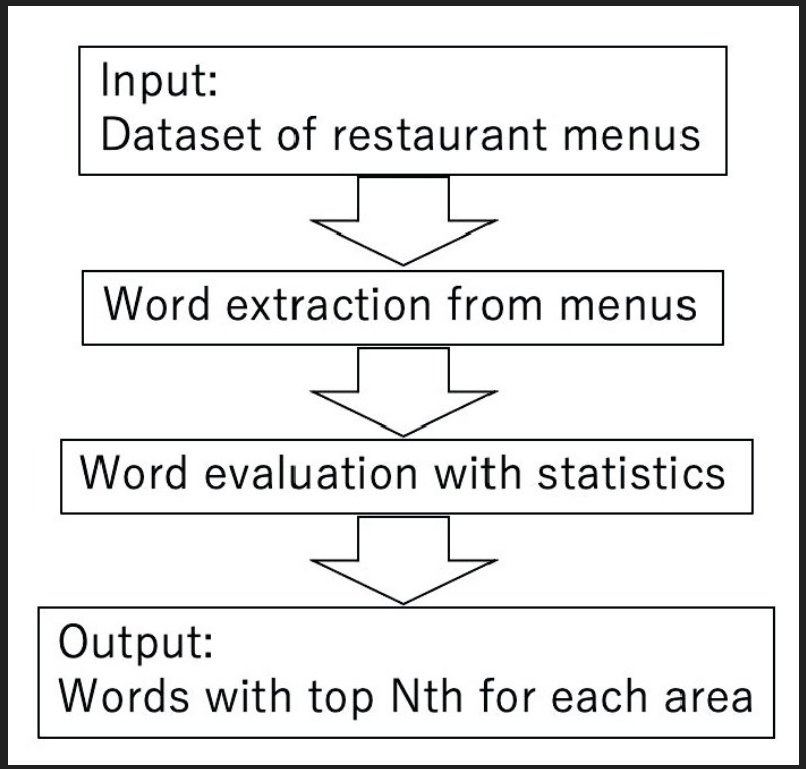
Contributions of this paper

(1) We developed a dataset of restaurant menus in Japanese roadside stations. The dataset has 8,707 menus from 1,109 roadside stations in 47 prefectures (regions).

(2) We discovered that the keywords for the local foods were not only foods but also place names, dish names, and their combinations.

Outline of the proposed method

- Figure shows the outline of the proposed method.



Dataset of restaurant menus in roadside stations

- The roadside stations have Web pages that show restaurant menus.
- We used the following attributes:
  (1) Restaurant menu.
  (2) Roadside station name of the restaurant offering the menu.
  (3) Prefecture, where the roadside station is located.
- We choose prefectures in Japan as the regions.

TABLE II: STATISTICAL DATA OF DATASET.

Item	Value
Number of prefectures	47
Number of roadside stations	1,109
Number of menus	8,707
Average of menus per a prefecture	184.8
Minimum of menus among prefectures	0
Maximum of menus among prefectures	561
Average of menus per a roadside station	24.9
Minimum of menus among roadside stations	0
Maximum of menus among roadside stations	102

Examples in the dataset

Menu	Roadside station	Prefecture
Chinese noodle with horse meat	Shichinohe	Aomori
Chinese noodle with boiled strawberry	Hashikami	Aomori
Curry with Murakami (place name) beef	Asahi	Nigata
Rice with tea poured and a sea bream in Kashiwazaki (place name)	Nigata furusato mura	Nigata
Meal of fried pork with Japanese miso	Fuji	Shizuoka
Japanese mustard flavored ice cream	Amagi goe	Shizuoka
Meal of fried pork of Fujisakura (place name)	Mitomi	Yamanashi
Row salmon of Shinshu (place name)	Kotani	Yamanashi
Meal of fried chicken of Choshu (place name) with vinegar and tartar sauce	Abucho	Yamaguchi
Meal of rice bowl with fried shrimp of Hagi (place name) and Japanese noodle	Hagi jyukan	Yamaguchi
Cold Japanese soba of Reiwa (era name of Japan's official calendar)	Adachi	Fukushima
Meal of Aizu (place name) DE Jyaran	Kita no sato	Fukushima
Rice bowl of fried gamecock with vinegar and tartar sauce	Nangoku furari	Kochi
Rice bowl of shiitake mushroom	Birafu	Kochi
Ice cream with fig flavored	Buzen okoshikake	Fukuoka
Ice cream with salt and Yame (place name) tea flavored	Tachibana	Fukuoka
Curry with Japanese pepper	Shimizu	Wakayama
Fried chicken	San Pin Nakatsu	Wakayama
Curry with loquat	Tomiura	Chiba
Meal of preserved fish	Kamogawa ocean park	Chiba

Keyword extraction for a local food

- Keyword extraction is performed with a statistical index TF-IDF.
  - TF-IDF is an index of word importance.
  - TF: word frequency, IDF: inverse document frequency
- It is assumed that the **keywords for the local foods frequently appear in the menus of the area (RF)**. In contrast, **the keywords for the local foods do not frequently appear in those from other areas (IRF)**.

$$RF - ILF(w, a) = rf(w, a) \times log \frac{M + 1}{lf(w, A) + 1},$$

w: word, a: prefecture,
M: number of prefecture
A: set of prefecture

  - rf: restaurant frequency, lf: local frequency

Experiment

- Procedures
  - (1) An experimenter extracted keywords using the proposed method.
    - The experimenter extracted 20 keywords for each of the 47 prefectures in Japan, totaling 940 keywords.
  - (2) The experimenter asked participants to evaluate whether or not the extracted keywords are for the local foods.
    - 92 Japanese (45 male and 47 female) participated in the experiment.
  - (3) The experimenter evaluated the rate of keywords for the local foods.
    - We obtained 100 answers for 29 prefectures.

Experimental results

- We obtained answers from 29 of the 47 prefectures. The highest, lowest, and average approval rates were 68%, 4%, and 21.1%, respectively.
- The approval rate was the rate of approval numbers versus answer numbers.
- Table (→) shows rates for local foods in top and bottom five prefectures.

Prefecture	Rate
Hokkaido	68%
Nagasaki	64%
Okinawa	39%
Gifu	30%
Aomori	28%
Hiroshima	4%
Ibaragi	5%
Miyazaki	6%
Kanagawa	8%
Yamaguchi	10%
Average	21.1%

Discussion: effectiveness of the proposed method

- Table (→) shows the extracted keywords and approval rates by the participants. The prefecure is Hokkaido.
- More than half of the participants approved 14 of the 20 keywords.
  - However, some keywords were judged as not the local foods (for example, curry and ice cream.)

Keyword	Rate	Type
Siretoko	100%	Place name
Rausu	66%	Place name
Chinese noodle	100%	Dish name
Ice cream	33%	Dish name
Curry	0%	Dish name
Oyster	100%	Food name
Scallop	66%	Food name
Fried noodles with starchy sauce	33%	Dish name
Hanamaka	100%	Place name
Kelp	100%	Food name
Crab	66%	Food name
Traditional Fried Chicken	100%	Dish name
Dosan	100%	Place name
Lunch	33%	Others (meal style)
Salmon roe	100%	Food name
Mongolian mutton barbecue	100%	Dish name
Sun flower	33%	Others (plant name)
Atsukeshi	0%	Place name
Uryu	66%	Place name
Vegetable	66%	Food name

Discussion: types of keywords extracted by the proposed method (1)

- The extracted keywords are for food and non-food items.
  - place names, dish names, meal styles, and plant names.

Keyword	Rate	Type
Siretoko	100%	Place name
Rausu	66%	Place name
Chinese noodle	100%	Dish name
Ice cream	33%	Dish name
Curry	0%	Dish name
Oyster	100%	Food name
Scallop	66%	Food name
Fried noodles with starchy sauce	33%	Dish name
Hanamaka	100%	Place name
Kelp	100%	Food name
Crab	66%	Food name
Traditional Fried Chicken	100%	Dish name
Dosan	100%	Place name
Lunch	33%	Others (meal style)
Salmon roe	100%	Food name
Mongolian mutton barbecue	100%	Dish name
Sun flower	33%	Others (plant name)
Atsukeshi	0%	Place name
Uryu	66%	Place name
Vegetable	66%	Food name

Discussion: types of keywords extracted by the proposed method (2)

- Table (→) shows types of keywords extracted by the proposed method.
- Words with the highest score in a prefecture were analyzed.
- The results indicated that the keywords for the local foods were not necessarily foods.
- People could feel the locality from keywords for place and dish names.

Type	Rate (total 47 prefectures)
Place name	40%
Dish name	26%
Food name	17%
Combination of place and food names	11%
Person name	4%
Adjective	2%

Conclusions

- This study applied a basic statistical method for extracting keywords for local foods from restaurant menus of roadside stations.
- Contributions of this paper are summarized as follows:
  - (1) We developed a dataset of restaurant menus in Japanese roadside stations. The dataset has 8,707 menus from 1,109 roadside stations in 47 prefectures (regions).
  - (2) We discovered that the keywords for the local foods were not only foods but also place names, dish names, and their combinations.
- In the future, we will develop a map for the local food restaurants to promote tourism.