International Journal of u- and e- Service, Science and Technology
Vol.9, No. 1 (2016), pp.39-52
http://dx.doi.org/10.14257/ijunesst.2016.9.1.05

Food Recommendation System Using Big Data Based on Scoring Taste Adjectives

Suk-Jin Kim[1], Hee-Jue Eun[2] and Yong-Sung Kim[3]

Department of Computer Engineering, Chonbuk National University 664-14 DuckJinDong, DuckJin-Gu, Jeonju, Korea
hanuri00@jbnu.ac.kr[1], hjeun@jbnu.ac.kr[2]
Corresponding Author, Yong-Sung Kim, yskim@jbnu.ac.kr

Abstract

It is a general practice to evaluate food taste based on sensory tests, however, this test method's disadvantage is that a lot of cost and time is required and significant deviation is taken place depending on each evaluator as well. Food taste evaluation by utilizing SNS-based big data for supplementing this disadvantage is considered to be a new challenge and innovative method. The objective of this study is to suggest a system that evaluates and recommends the level of domestic food taste by not only clustering food preference using k-means algorithm after sorting out food-related tweet contents from typical twitter of SNS, followed by scoring taste adjectives being mainly used in daily life by using rough set and selecting food-related adjectives among the scored adjectives, but also exploring the level of salty, sour, savory, bitter and sweet tastes through perception map.

Keywords*: Food taste, Taste Adjective, Recommendation System, k-means Algorithm, Scoring taste adjectives, clustering, Sensory evaluation, Bigdata, twitter*

1. Introduction

At an early stage, a research on sensibility measurement technology of olfactory/gustative sense for food taste and DB development was performed by MOE in 1998, however, subsequent research has seldom been made. Taste depends not only on five tastes (salty, sour, sweet, bitter and umami tastes) and spicy taste but also on visual, olfactory effect and texture, and warmth of food. However, so far, a general analysis of taste adjectives being used by contemporary people as a basis for researching on taste has been hardly performed in our country [1,2].

Research on taste expression of Korean food is mostly performed through sensory evaluation. Sensory evaluation is defined as an area of science that measures, analyzes and interprets a reaction of food and substances features being felt by 5 senses (visual, olfactory, gustative, tactile, and auditory senses) [3]. In this study, in order to score taste adjectives, an integrated analysis was performed through a questionnaire by having around 30 sensory evaluation team attending food, nutrition department, and food engineering department taste, evaluate food taste and taste adjectives for the sensibility of those team recognizing and sensing such food were extracted.

Around 20,000 samples of taste adjectives relevant to all the tastes including gustative, olfactory senses, texture, warmth being used in our country were extracted from foodservice management magazine (from January, 2014 version to October, 2014 version). In order to score the conformity of collected taste expression adjectives, 108 taste expression adjectives were selected by composing taste-scoring table through correlation analysis and multivariate analysis that are statistic method.

ISSN: 2005-4246 IJUNESST

71 adjectives were finally selected after comparatively analyzing taste adjectives being extracted like this with 87 taste adjectives data including gustative sense, texture, warmth and olfactory sense, and through rough set theory, taste adjectives were scored. By using these 71 scored adjectives, preferred foods among that favored by Koreans were sorted out through SNS big data-based k-means algorithm and features of sorted specific food tastes were perceptualized. Through this, features of Korean taste were analyzed and based on this analysis, a recipe method of food was established depending on features of foreigner's taste.

2. Relevant Research

2.1. Big Data and SNS

2.2.1. Big Data

Research utilizing big data is being actively performed in each field. When observing research relevant to big data, it is mainly utilized in traffic, health, disaster prevention, politics, economy, culture and tourism but in food-related field, research has never been attempted. In other words, as shown on [Fig. 2-1] WordCloud, which was prepared by keyword of app. 800 pieces of theses from 1990 to 2015 as big data, there was no research that took food as its topic.

Figure 2-1. Thesis Keyword Cloud

2.2.2. SNS (Social Network Service)

SNS is two-way communication means that facilitates interaction among users based on an open type approach method (Kozinets *et al*., 2010). It is also an online tool and media platform that enables information accumulation through open style information sharing and communication [13]. This is a service focused on formation of online social relations for the community who wishes to observe other people's personal taste and activities [14]. When observing a research on food and taste by utilizing SNS, a research on a system for automatically recommending food, beverage coupon in social commerce suitable for the users considering situational information such as purchase history [17], interaction of SNS for accommodation intention of foodservice-related information [18], a research on data cooking device and service that helps implement cooking handily through seasoning mixture by providing diversified recipes reflecting needs of single consumers[19] and a research that utilizes SNS in marketing analysis by identifying consumption behavior features of domestic foodservice consumers after classifying foodservice consumers and SNS utilizing types [20] are available.

2.2. Taste Adjectives

Humans perceive stimulants of the outside environment through 5 senses including visual, auditory, olfactory, tactile and gustative sense. Among these, gustative sense

recognizes salty, sour, sweet, and bitter tastes through chemical reaction of tongue. [Fig. 2-1] shows an expression of taste adjective words.

Figure 2-2. Taste Adjective Wordfield

In language, words are used having a relation with other words, and are not used independently. This relation could represent connectivity between two words or among groups of words. Therefore, connectivity of words could be expressed as a word set having similar meaning, which is then called wordfield.

At an early stage, Tina Nabrosky tried to structure adjectives of taste expression as wordfield. As shown on [Fig. 2-3], he said that taste quality, intensity, and pleasure that are three parameters in classifying taste adjectives like 3D wordfield play a decisive role [22].

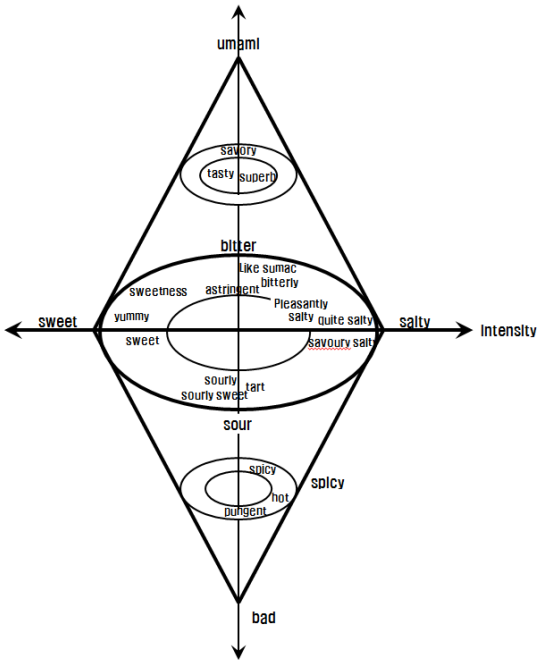

Figure 2-3. Wordfield of Taste 3D

Generally, taste quality is defined as 5 tastes including sweet, sour, salty, bitter, and umami tastes, however, it was also classified as 6 tastes including spicy taste that was added to Korean food taste as an essential element.

2.3. Rough Set Theory and Clustering

2.3.1. Rough-set Theory

Rough set is a mathematical approach suggested by Pawlak [25] and it could be applied to diversified fields including automatic classification, pattern recognition, learning algorithm [26].

As research using rough set, a research that suggested an algorithm extracting features comprising sound source of musical instrument by using entropy concept of rough set for the inter-dependency level of game music characteristics [27], a research suggesting a design method of gas identification system having hierarchical rule using rough set [28] and a research using rough set for selecting set of important characteristics among the entire characteristics in suggesting fuzzy discretization technique for the data having continuous type attribute [29]are available. Lang and Nguehn suggested tolerance relation-based rough clustering in 2002 and it is regarded as a successful method in document clustering. When comparing document clustering to our food clustering, document and terms in document clustering may be called as adjective relevant to food twit and taste in this report.

2.3.2. Cluster Analysis

Cluster analysis is a technique that is used in analyzing data by dividing it into clusters sharing specific nature. In addition, as a clustering technique, it is classified as hierarchical clustering, divided clustering, neural network, and statistical search technique, and when observing it in detail, first, as hierarchical clustering technique, there is a dendrogram that represents clustering results as a hierarchy and second, as divided clustering technique, an arithmetic operation that assigns each sample to cluster is used. Third, as neural network, there is SOM (Self-organizing map), ART (Adaptive Resonance Theory) technique and fourth, as a statistical search, SA (Simulated Annealing) technique, genetic algorithm could be cited [30,31,32]. Cluster analysis is a technique being used in analyzing data in the fields of taxonomy, information retrieval, climatology, psychology, medical science, and business. A hybrid method that mixes data properly by utilizing various cluster techniques is reported in the latest thesis. More than anything else, cluster method and number (count) fit for cluster features are important, and if clustered clusters should be well expressed visually and classified, it would be helpful for understanding cluster.

3. Recommendation System Design based on Taste Adjectives

3.1. Outline of Recommendation System

A recommendation system provides propensity and preference such as recommendation of goods, maker brand based on price ,and performance in ranking order and nowadays, it is also utilized in information retrieval field like the page rank algorithm of Google. The food recommendation system suggested in this study is performed based on the following procedure.

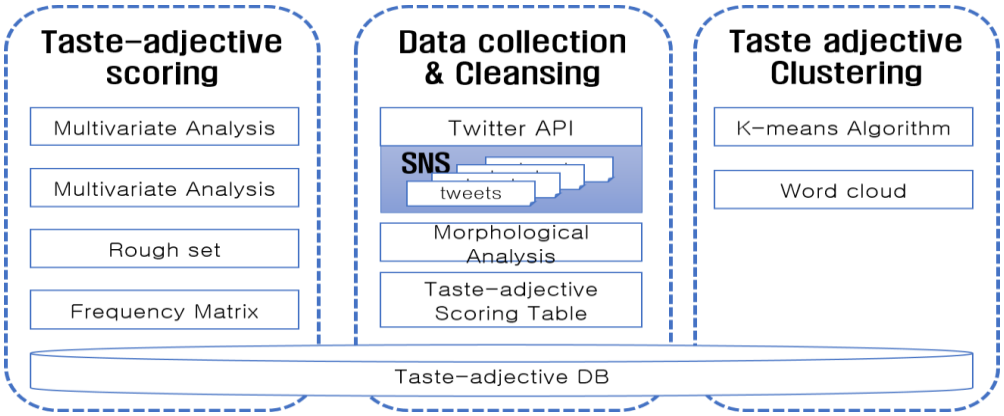

Figure 3-1. Taste Adjective-based Clustering Procedure

(1) Preparation of Taste Adjective Scoring Table

In order to score conformity of taste adjectives after extracting approximately 20,000 samples of taste adjective from foodservice magazines (from January, 2014 version to October version) relevant to major taste adjectives for Korean language expression, correlation and multivariate analysis for taste adjectives were performed and through this correlation, 108 taste adjectives were extracted. By comparatively analyzing extracted taste adjectives with 87 taste adjective data for gustative sense, texture, temperature and olfactory sense, 71 taste adjectives were sorted out and were then scored through a rough set of such adjectives.

If comparing document clustering to the clustering of our food, in document clustering, document and terms are food twit, taste adjective and its frequency matrix ($AF=[at_{i,j}]_{NxM}$) in this study and where, $at_{i,j}$ is j numbers of taste adjective expressed in ith food and N is expressed as number of twit and M as number of taste adjective. In this study, N is 21038 ea and M is 71 ea, respectively.

$$T = (U, A \cup d)_{T=(U,Ad)}$$

U : tweets

A : $\alpha : U \rightarrow V_{\alpha}$ forever $\alpha \in A$ evaluation scale

V_a : Value of evaluation scale

$\alpha \in A$: Taste adjective

Table 3-1. Taste Adjectives

bitter		sour		salt		sweet		umami		spicy		taste status	
taste adjective	score	taste adjective	score	taste adjective	score	taste adjective	score	taste adjective	score	taste adjective	score	taste adjective	score
like sumac	0.8	sour	0.75	salty	0.82	Honey-like	0.86	savory	0.82	hot	0.86	tasty	0.99
bitter	0.78	Tart	0.71	quite salty	0.78	sweet	0.72	tasty	0.82	spicy	0.69	fantastic	0.76
bitterly	0.72	sourly	0.65	slightly salty	0.64	Scorched smelling	0.71	delicious	0.80	thirsty	0.54	good	0.73
quite bitter	0.71	sourly sweet	0.62	Savoury salty	0.63	Sweetish	0.63	appetizing	0.79	pungent	0.61	nothing	0.32
Astringent	0.67	refreshing	0.57	Pleasantly salty	0.34	sweetness	0.54	superb	0.69	burning	0.58	tasteless	0.13

The algorithm of scoring, modifying and supplementing taste adjectives and preparing its scoring table is as shown on below [Fig 3-2].

Algorithm 3.1 Taste-adjective scoring

```
input : terms
output : taste-adjective tweets
Begin
        collecting taste-adjective from sample data          //dictionary, magazine, etc.
        selecting and averaging taste-adjective using corelation and
        multivariant analysis
        for (i=1 to N {                                      // N of adjective-terms
                webinar() ;                                  // rough Set using R package
                frequency();                                 // taste strong rate
        }
        display() ;                                          taste-adjective table
End
```

Figure 3-2. Taste-adjective Scoring

(2) Data Collection

In this study, for the data to be utilized at the stage of data analysis, Twitter being most frequently used among social networks is used and the procedure is as follows.

First, collect approximately 300,00 tweets of tweeters by using Twitter API from January 2015 to August 2015. Second, collected tweets are processed as stopwords through morphological analysis process and only relevant tweets are stored in DB by using a taste adjective scoring table being generated in a process of pre-processing. Below [Fig 3-3] is pseudo-code of system constructing taste-adjective DB through stopword processing.

Algorithm 3.2 built raw-taste-DB

```
input : tweets
output : taste-DB
Morphological analysis and treatment as stopwords
Begin
        collecting tweets;                                   // using Twitter API
        for (i=1; i<N ; I++{                                 // N of tweets
                if (tweets = taste-adjective table){         // using table-element
                        copy taste-DB ;                      // No of accumulate
                        cnt = cnt +1 ;
        }
        display()                                            // Not of tweets(cnt)
End
```

Figure 3-3. Built Raw-taste-DB

(3) Data Analysis

The stage of data analysis is an analysis process of recommending foods by using collected data and its procedure is as follows.

Algorithm 3.3 Data Filtering

```
input : tweet data
output : cleaned Noun-data
Begin
        reading tweets data                                  // data in which stopwords are removed
        save rwadata -> txt file encoding UTF-8
        load useSejongDic()                                  // Additional pre-construct of taste adjectives
        insert into useSejongDic() value new taste-adjective
        twitter data filtering using on useSejongDic()       //noun extraction
End
```

Figure 3-4. Sata Filtering

Below [Fig 3-5] shows a clustering process by using k-means algorithm utilizing cleaned noun-data that is finally generated through removal of stopwords from source (raw) tweet, taste adjective scoring, DB and Sejong Dictionary.

Algorithm 3.4 k-means clustering

```
input : cleaned Noun-data
output : clustering map
Begin
        initialize Z()                          // initialize the Z of number of k
        while(TRUE) {                           //
                for (i=1 to N)                  // x_i position at the nearest clusters center
                if (old_Z = new_Z) break;       // compare with previous cluster center
                for (j=1 to k)                  // Z_j replace Z_j
        }
End
```

Figure 3-5. k-means Clustering

4. Implementation

4.1. Source (Raw) Data

In this study, Twitter data was collected from January 1 to August 31 by using Twitter API. Contents of importing collected approximately 300,000 raw data to memo screen is as shown on below [Fig. 4-1].

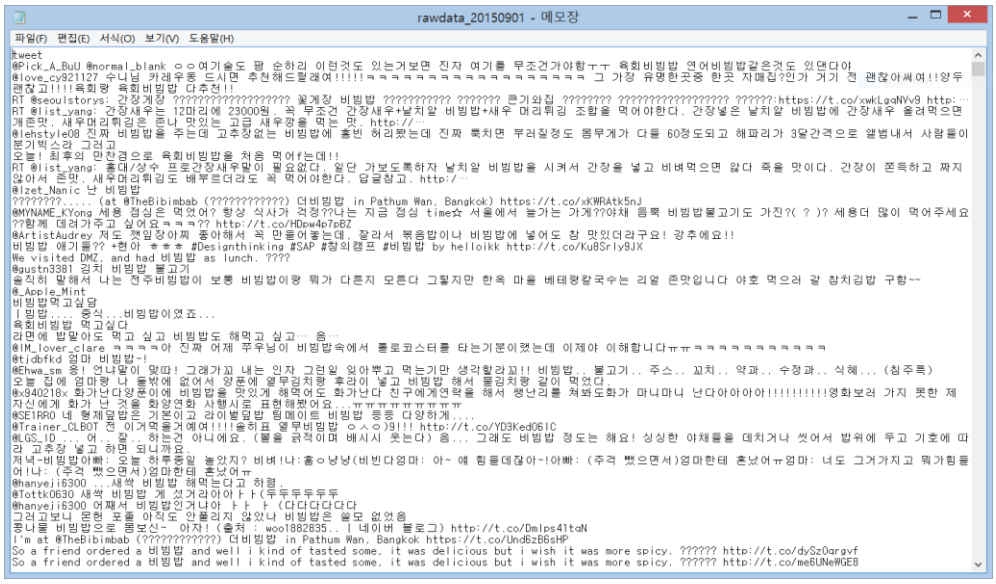

Figure 4-1. Raw Data

Among these approximately 300,000 tweets, 45076 cases of significant tweets were extracted after primarily removing insignificant tweets having nothing to do with taste adjectives and tweets that do not include taste adjectives. Afterwards, in order to ensure data simplicity and statistic processing, tweets that repeated the tweet of same content for more than 2 times and posted same contents of same person for more than 2 times were removed. In this study where repeated tweets were removed, number of valid tweets was reduced to 21038 cases.

4.2. Morphologic Element Analysis

Morphologic analysis for raw data was processed by utilizing the KoNLP package of R. As taste-related adjectives in addition to proper nouns were not included in Sejong Dictionary included in KoNLP [42], only the root of taste adjectives and words were extracted after adding the scored taste adjective table separately. Below [Fig. 4-2] randomly shows 500 words for the data after extracting only root of taste adjective and word.

Figure 4-6. Extraction of Word Root

Twitter data having taste adjectives was filtered from root of extracted taste adjectives and words by using the taste adjective scoring table. Approximately 21038 valid tweets data in which tweets doubled for more than 2 times or tweets comprising only insignificant stopwords were removed from total 45076 data were filtered after stopword processing and morphologic analysis. Among these, top ranking words having highest frequency are as shown on below [Fig. 4-3].

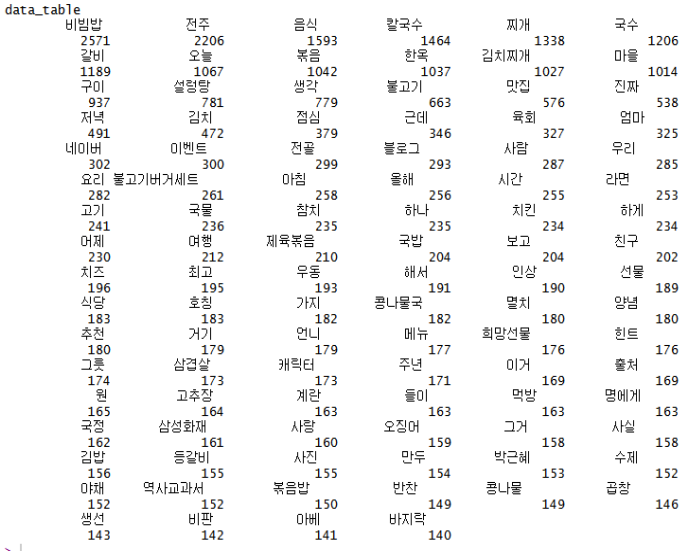

Figure 4-3. 100 Words Having High Frequency Ranking

By using words having a frequency of 2 times among taste adjectives and words in 21038 tweets, a wordcloud was comprised. Below [Fig. 4-4] shows wordcloud composition through R. Bibimbap was most frequently mentioned and Jeonju, food, Kalguksoo (handmade noodle soup) and Korean stew were also mentioned occasionally.

Figure 4-4. Food and Taste Adjective Words

4.3. Clustering

After processing filtered data based on 5 tastes including sweet, salty, sour, bitter, umami taste, clustering was performed. Distribution of high ranking 20 words mentioning food is as shown on below Figure. Most frequently mentioned word was 'Bibimbap' as 11.7%, followed by Jeonju 10.0%, food at 7.2%, Kalguksoo at 6.7%, and Korean stew 6.1%. The reason why Bibimbap was frequently mentioned was attributable to diversity of Bibimbap [8] and an analysis of integrating Yukhoi-Bibimbap, Nutrient Hot-pot rice-Bibimbap, Dolsot-Bibimbap, Sprout -Bibimbap, Wild vegetable-Bibimbap as just Bibimbap. [Table 4-9] shows analysis processing value based on scoring table for 6 kinds of food being most frequently mentioned in collected tweeters. Typical classification of taste for Bibimbap was considered to be spicy food and its spicy level was represented to be the highest as 0.78 followed by sweet taste (0.68), sour taste (0.59), umami taste (0.53), salty taste (0.43) and bitter taste (0.42), respectively. Umami taste of Bulgogi was represented to be high as 0.85.

Taste perception map is as shown on below [Fig. 4-5] and [Fig 4-6].

Taste	umani	sweet	sour	spicy	bitter	salty
Bibimbap	0.53	0.68	0.59	0.78	0.42	0.43
Bulgogi	0.85	0.74	0.30	0.48	0.37	0.45
Gabitang	0.78	0.67	0.41	0.34	0.43	0.65
Seolreongtang	0.82	0.73	0.37	0.27	0.38	0.53
Kalguksoo	0.56	0.54	0.31	0.53	0.27	0.47
Kimchi-stew	0.54	0.43	0.24	0.78	0.55	0.60

Table 4-1. Taste Perception Map for Foods (6 Kinds)

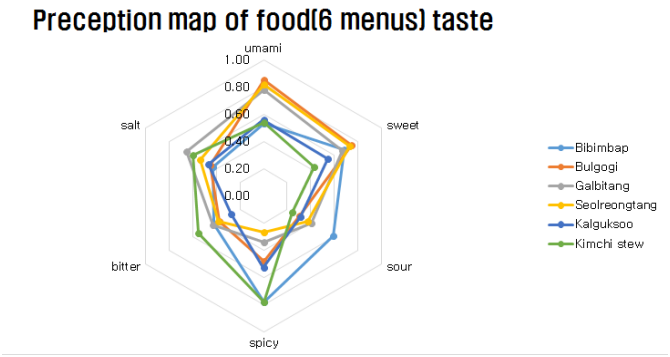

Figure 4-5. Perception Map of Food (6 Menus) Taste

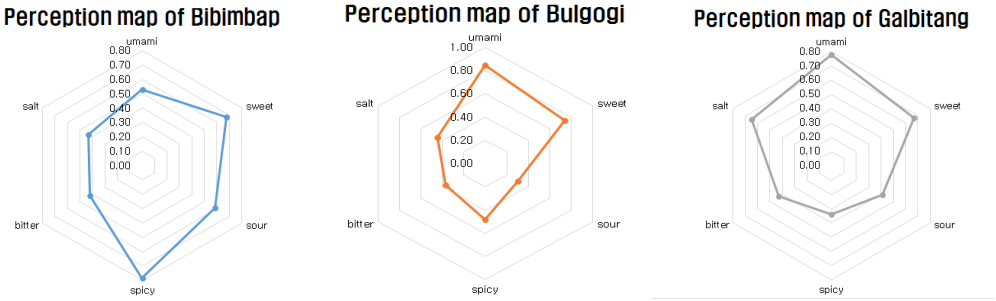

Figure 4-6. Perception Map of Popular Korean 3 Menu

Taste adjective clustering diagram is as shown on [Fig. 4-12]. Considering taste adjectives of 6 classifications including 5 tastes and spicy taste, the number of initial clustering was defined as 6. By limited taste adjectives, clustering distribution shows lineal shape. It is required to be analyzed through more augmented taste adjective and tweet data

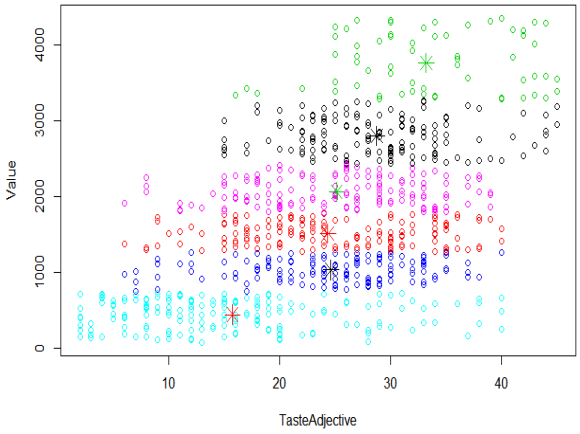

Figure 4-6. Taste Adjective Clustering Perception Map

5. Conclusion

As the advanced countries are doing their effort of globalizing their own food earlier as a part of globalizing their cultures, American hamburger, Japanese sushi, Chinese dimsum, Italian pasta and pizza, French tart and wine, Swiss fondue, Indian curry, Vietnamese rice noodle, Turkish kebab, and U.K.'s fish and chips have now become global foods.

In this study, the following procedure was performed by starting a research based on ideas that globalize Korean traditional food, what is the most favorite food of Koreans, why such food became food preferred by Koreans, why foreigners still hesitate to prefer this food and sensibility of feeling taste differs by each country and age.

In order to sort out taste adjectives expressing taste first, taste adjective words being utilized at present were extracted and sorted out again through sensory evaluation and then taste-adjective table was composed by scoring sorted taste adjectives through correlation, multi-variate analysis and rough set fit for 5 tastes.

And inIn order to select Korean typical foods, data was collected by utilizing tweeter Twitter among social networks and tweets belonged belonging to tthe aste-adjective table were sorted out through morphologic analysis and table mapping. and then Bibimbap aAmong Korean style food having high ranking frequency, Bibimbap was selected, and in particular,with Yukhoi-Bibimbap in particular as the higher preferred among various Bibimbap being sorted out by k-means algorithm and recommended.

In addition, by showing a perception map for 5 tastes in order to measure the taste intensity of Yukhoi-Bibimbap, the taste sensibility of Yukhoi-Bibimbap that is typical Korean food could be obtained.

Contribution of this study is to score Korean taste-adjectives for the first time and suggest algorithms through which typical Korean food could be recommended by using scored taste-adjectives and provide a foundation for utilizing this algorithm.

Future research task is to suggest tastes and recipes that may globalize Korean food through preparing a taste-sensibility map of each city and province, and the taste map of foreigners by broadly implementing regionally collected Twitter data to nationwide level.

References

[1] C. Kim, “A Study on Meaning of the Taste Adjective”, The Korean Language and Literature 100, vol. 6, **(2008)**, pp. 1-30.
[2] S. Lee, “A Study on the Structure of Korean Taste Terms”, The Korean Language and Literature 100, vol. 2, **(1988)**, pp. 33-38.
[3] “Sensory test”, Institute of food Technologists (IFT).
[4] “Identify research excellence of Korean food”, Ministry for Food, Agriculture, Forestry and Fisheries (MIFAFF), **(2008)**.
[5] “Korean Ministry for Food”, Agriculture, Forestry and Fisheries, **(2010)**.
[6] Y. Eomand J. Kim, “An Analysis of the Effects of Jeonju Bibimbap Price-quality Attributes Association on Tourists’ Satisfaction and Behavioral Intentions”, Tourism Study, vol. 30, no. 3, **(2015)**, pp. 201-222.
[7] Y. Joo, “A Study on Evolution and Discourse of BIBIMBAP”, Society and history 87, vol. 9, **(2010)**, pp. 5-38.
[8] Y. Eom and J. Kim, “An Analysis of the Effects of Jeonju Bibimbap Price-quality Attributes Association on Tourists’ Satisfaction and Behavioral Intentions”, Tourism Study, vol. 30, no. 3, **(2015)**, pp. 201-222.
[9] Gartner, “How to Plan, Participate and Prosper in the Data Economy”, Gartner Group, vol. 1, **(2011)**.
[10] I DC, Digital Universe study, **(2011)**.
[11] McKinsey & Company, “Big data: The next frontier for innovation, competition, and productivity”, **(2011)**.
[12] “Related Articles Keyword Analysis Big Data, Papers Big Data between 1990 and 2015”, National Research Foundation of Korea, **(2015)**.
[13] Kaplan and Haenlein, **(2010)**.
[14] Report by IWGDPT (International Working Group International Working Group on Data Protection in Telecommunications), **(2008)**.
[15] “Internet usage”, KISA (Korea Internet & Security Agency), vol. 12, **(2014)**, pp. 23-35.
[16] Report by Koea Information Sciety Development Institute, **(2013)**.

[17] S. Y. Lee and H. J. Jeon, "Food coupons Recommendation System using Bayesian Network for Social Networking Service", Proceedings of KIIS Fall Conference, **(2012)**.
[18] S. Kim and S. Jun, "The Effects of SNS Users Characteristics and Informational alues on the Information Acceptance Intention", The Journal of Internet Electronic Commerce Resarch, vol. 14, no. 6, **(2014)**, pp.127-146.
[19] B. Kim, E. Lee, D. Kim, I. Lim, I. Yu and S. Y. Ki, "Proposal for an SNS-based sauce-miing device for improvement of single consumers cooking environment Centering on a behavior survey of single people", Extended Abstracts of HCI KOREA, vol. 12, **(2014)**, pp.225-229.
[20] H. Sung and S. Lee, "SNS Usage - based Segmentation of Restaurant Customers", Journal of Tourism and Leisure Research, vol. 26, no. 4, **(2014)**, pp. 189-204.
[21] H. M. Kim, "Comparative analysis of multi-cultural situations between Korea and Germany", Proceedings of the International Workshops on FRS, **(2013)**; Jeju, Republic of Korea.
[22] H. J. Buckenhüskes, K. von Heusinger and T. Nawrocki, "Die Dimensionen des Geschmacks", Bedeutungsfelder für Geschmacksadjektive, **(2004)**.
[23] Chonbuk National University Institute for Intercultural Communication, "Taste Adjective Dictionaly for the Globalization of Korean Food", Chonbuk National University Press Publishers, ChonBuk National University, **(2013)**.
[24] J. Lee, K. Park and J. Rh, "Fuy Relation-Based Analysis of Korean Foods and Adectives for Taste Evaluation", Journal of Korean Institute of Intelligent Systems, vol. 23, no. 5, **(2013)**, pp. 451-459.
[25] Z. Pawlak, "Rough sets", INternational Journal of Computer and Information Sciences, vol. 11, **(1982)**, pp. 341-356.
[26] Z. Pawlak, "Rough Sets-Theoretical Aspects of Reasoning about Data", Kluwer Academic Publishers, **(1991)**.
[27] J. You and I. K. Par, "Musical Instruments Sounds Classification based on Rough Set Analysis", The Korean Society for Computer Game, vol. 26, no. 4, **(2013)**, pp. 67-73.
[28] Y. K. Bang, H. Byun and C. H. Lee, The Transaction of the Korean Institute of Electrical ENINEERS, vol. 61, no. 8, **(2012)**, pp. 1164-1171.
[29] C. S. Son, A. Shin, I. H. Lee, H. J. Park, H. S. Park and Y. N. Ki, "Fuy discretiation with spatial distribution of data and Its application to feature selection", Korean Institute of Intelligent Systems, vol. 20, no. 2, **(2010)**, pp. 165-172.
[30] N. Y. Su, B. M. Stith, H. Puche and P. Bardunias, "Characterization of Tunneling Geometry of Subterranean Termites (Isoptera: Rhinotermitidae) by Computer Simulation", Sociobiology, vol. 44, no. 3, **(2004)**, pp. 471– 483.
[31] S. H. Lee, N. Y. Su and P. Bardunias, "Optimal Length Distribution of Termite Tunnel Branches for Efficient Food Search and Resource Trans-portation", Biosystems, vol. 90, **(2007)**, pp. 802– 807.
[32] Oracle, Oracle : Big Data for te Enterprise, Oracle White Paper, Http://www.oracle.com/us/products /database/ big-data-for-enterprise-519135.pdf, **(2012)**.
[33] S. G. Cho and S. B. Ki, "Finding Meaningful Pattern of Key Words in IIE Transactions Using Text Mining", Journal of the Korean Institute of Industrial Engineers, vol. 38, no. 1, **(2012)**, pp. 67-73.
[34] Y. A. Sung, "Cluster Analysis for the Consumption Expenditure Patterns of One-Person Households of Different Age Groups", Korea Consumer Association, vol. 24, no. 3, **(2013)**, pp. 157-181.
[35] S. Park, Y. Lee, C. Shim and S. R. Le, "Generic Document Summariation using Coherence of Sentence Cluster and Semantic Feature", The Korea Institute of Information and Communication Engineering, vol. 16, no. 12, **(2012)**, pp. 2607-2613.
[36] K. Kim, "Dynamic Programming-Based Multiple Dim Small Target Detection Using Kmeans Clustering Algorithm", The Korean Society for Aeronautical and Space Sciences, **(2010)**, pp. 643-646.
[37] P. Maheswary and N. Srivastava, "Retrieval of remote sensing images using colour & texture attribute", International Journal of Computer Science and Information Security, vol. 4, no. 1, **(2009)**, pp. 19-23.
[38] P. S. Hiremath and J. Pujari, "Content based image retrieval using color, texture and shape features", Proc. of the 15th International Conference on Advanced Computing and Communications (ADCOM), Article (CrossRef Link), **(2007)**.
[39] W. Raffe, J. Hu, F. Zambetta and K. Xi, "A dual-layer clustering scheme for real-time identification of plagiarized massive multiplayer games (MMG) assets", Proc. of the 5th IEEE Conference on Industrial Electronics and Applications (ICIEA), **(2010)**; Taiwan.
[40] H. Jeon, KoNLP: Korean NLP Package. R package version 0.76.9. http://CRAN.R-project.org/ package=KoNLP, **(2013)**.

