Adaptive Method for Trends in Ranking of Tourist Spots

Yusaku Takano
Graduate School of System Design
Tokyo Metropolitan University
Hino, Tokyo, Japan
Email: takano-yuusaku@ed.tmu.ac.jp

Masaharu Hirota
Department of Information Science
Okayama University of Science
Okayama, Japan
Email: hirota@mis.ous.ac.jp

Daiju Kato
WingArk First Inc.
Shibuya, Tokyo, Japan
Email: kato.d@wingarc.com

Tetsuya Araki
Graduate School of System Design
Tokyo Metropolitan University
Hino, Tokyo, Japan
Email: araki@tmu.ac.jp

Masaki Endo
Polytechnic University
Kodaira, Tokyo, Japan
Email: endou@uitec.ac.jp

Hiroshi Ishikawa
Graduate School of System Design
Tokyo Metropolitan University
Hino, Tokyo, Japan
Email: ishikawa-hiroshi@tmu.ac.jp

***Abstract*—In recent years, for tourists deciding on a destination for a trip, demand for websites, such as TripAdvisor, for obtaining tourist information is increasing. These websites usually rank tourist spots with user reviews. Regarding tourist spots, trends appear in their popularity. When ranking tourist spots solely by reviews posted on these websites for obtaining tourist information, the possibility exists that new tourist spots and tourist spots with actively changing popularity are not ranked higher in rankings. However, a user needs to find the most enjoyable tourist spots at the time of the visit in the ranking of tourist spots. Therefore, we use tweets in this study as comments for tourist spots with high recency to generate a ranking of tourist spots considering popularity variation. Then, we model the ranking of tourist spots on TripAdvisor using tweets. After analyzing the contribution rates of classification for linguistic features and statistical features on tweets, we performed ranking of learning with effective features and generated a ranking of tourist spots considering popularity variation. Finally, our experimental results showed that tourist spots holding exhibitions or other activities frequently are ranked higher.**

Keywords–tourism; ranking learning; social recommendation.

I. INTRODUCTION

In recent years, with the increasingly widespread use of smartphones and tablets, people have become able to acquire useful information easily from the web. Along with this progress, when tourists decide on a trip destination, demand for websites, such as TripAdvisor [1], for obtaining tourist information is increasing. As the importance of the internet for travel increases, the way of using the internet during trips has changed [2]. In addition, particularly addressing user reviews, the importance of reviews on websites for obtaining trip-related tourist information has increased [3]. These websites usually present tourist spots in a ranking format and rank tourist spots with user reviews. Roughly speaking, the ranking of tourist spots is a comprehensive evaluation of past impressions. Additionally, these websites fundamentally keep the ranking algorithm secret to avoid deliberate manipulation of rankings and to prevent their use by other web services.

Regarding tourist spots, trends exist in the popularity of certain tourist spots. In other words, the popularity of tourist spots varies for different reasons, such as events or seasons. When ranking tourist spots using only reviews posted in the past, the possibility exists that new tourist spots and tourist spots with actively changing popularity are not ranked higher in among all rankings. However, most people might browse only higher ranked tourist spots. Users need to find the most enjoyable tourist spot at the time of their visit. In short, the ranking of tourist spots using only reviews posted in the past might not satisfy a user's need. For example, as an accurate ranking for a user's need, famous tourist spots for the viewing of cherry blossoms, such as Shinjuku Gyoen, can be expected to be ranked highly in the spring. For this study, we rank tourist spots in consideration of their current popularity variation in addition to the comprehensive evaluation of them in the past.

Many comments related to tourist spots are also posted to Twitter [4] and Facebook [5] which is a typical Social Network Service (SNS). These posts are comments about tourist spots with higher recency than reviews on websites for obtaining tourist information. In addition, these posts include user impressions about tourist spots. In this study, we propose an adaptive method for popularity variation of tourist spots using tweet messages from Twitter. Although we propose a method to make a ranking considering the latest popularity of tourist spots, the reliability of information related to user posts is low because many posts are made by heterogeneous users and robots on Twitter. Castillo et al. [6] pointed out a difficulty for the reliability of information related to posts. They proposed a method to extract only trusted information from Twitter. As described above, the information of tweets is unreliable. It is difficult to rank tourist spots solely by impressions of tourist spots given in the contents of tweets.

Therefore, in this study, our main idea of ranking tourist spots is modeling the ranking on TripAdvisor using tweets. Reviews by users in each post on websites for obtaining tourist information are more informative than Twitter and are reflected in the details for tourist spots. Although many websites exist for obtaining tourist information, TripAdvisor is a popular website for obtaining tourist information. Filieri et al. [7] examined the reliability and its influence by its contents and revealed that many tourists are using TripAdvisor. We will generate a ranking of tourist spots incorporating changes in popularity by application of recent tweets to the model. Consequently, tourist spots related to events and seasons will be ranked at the top of the ranking, which makes it possible to propose a ranking that satisfies the most enjoyable tourist spots at the time of the visit.

The remainder of the paper is organized as follows. Section II presents works related to feature selection for posts by users on the internet and tourism recommendations using posts on SNS. Section III presents a discussion of the results to analyze the contribution of classification with linguistic features and statistical features on tweets at tourist spots. Section IV presents results obtained using logistic regression and ranking learning with effective features for classification. Section V explains a discussion of the obtained results. Section VI concludes the paper with a discussion of results and avenues for future work.

II. Related Work

In this section, we present related works about analysis of user-posted reviews and recommendation of tourism spots.

A. Analysis of user-posted reviews

Currently, reviews on the internet are increasing. Many studies use these reviews. Mukherjee et al. [8] examined the spam filter system in Amazon Mechanical Turk with linguistic features and statistical features for reproducing the spam filter system in Yelp. In addition, Guy et al. [9] proposed a method of automatically collecting useful information for sightseeing with linguistic features on TripAdvisor. Therefore, in this study, we model the ranking of tourist spots on TripAdvisor by considering effective features with linguistic features and statistical features.

B. Recommendation of tourism spots

Impressions of tourist spots on SNS are increasing. These posts are comments by users with higher recency than reviews on websites for obtaining tourist information. Many studies applied these posts to recommendation systems for sightseeing. Borras et al. [10] classified these recommendation systems technically and considered their importance on trips. Ye et al. [11] [12] proposed a method to recommend geographical information, such as Points-Of-Interests (POIs), from check-in information and user attributes in Foursquare [13]. Choudhury et al. [14] and Lim et al. [15] proposed a method to recommend POI with user behavior on Flickr [16]. Ishihara et al. [17] reported a recommendation system for sightseeing while considering posts on Twitter from the perspective of sensitivity engineering. Mizutani et al. [18] proposed a recommendation system for sightseeing that is more suitable for users individually using posts on SNS, such as Twitter.

For these studies, they do not consider the effects of changes by enviromental factors and so on, but Missaoui et al. [19] pointed out the importance of taking into account changes of a user's preferences and environments for recommendation systems, and proposed a method to predict a user's preferences in the tourism domain. Additionally, these studies analyzed recommendation system for sightseeing only with posts and metadata on the SNS. However, many noise posts are sent to SNSs. The reliability of the information is questionable, as described in Section I. This study applies feature selection while considering the contribution rates of logistic regression. Also, this study proposes a system that generates a ranking of tourist spots incorporating a trend of tourist spots by modeling the ranking on the TripAdvisor where the reliability of information is higher than Twitter as training data.

In addition, various ranking methods use tweets. Duan et al. [20] proposed ranking recommendation to search tweets using ranking learning. Qupta et al. [21] proposed a recommendation method to rank important events with various linguistic features on tweets. Chang et al. [22] reported a method to remedy a shortage of recency on the ranking of web searches using tweets. This study models the ranking of tourist spots by application of ranking learning using tweets, which is one of the ranking methods.

III. Analysis of Features

As described herein, to examine the effective features for ranking learning, we divide the ranking of tourist spots into certain intervals and classified them. This section presents examination of the effective features for the classification of tourist spots and ranking learning from the distribution of each feature while considering linguistic features and statistical features of tweets.

A. Datasets

For this study, we use tourist spots among the top 100 rankings in Tokyo on TripAdvisor acquired in 2017. To analyze seasonal changes by month, we extracted tourist spots for which more than 100 tweets are posted each month. To obtain tweets at each tourist spot, we used GooglePlaceAPI [23] and NominatimAPI [24]. We were able to find tweets for a tourist spot if the latitude and longitude annotated to a tweet would be within the area of tourist spots obtained using these APIs. The number of tweets written in Japanese in 2017 was 1,691,521 at the 50 tourist spots.

B. Discussion of linguistic features

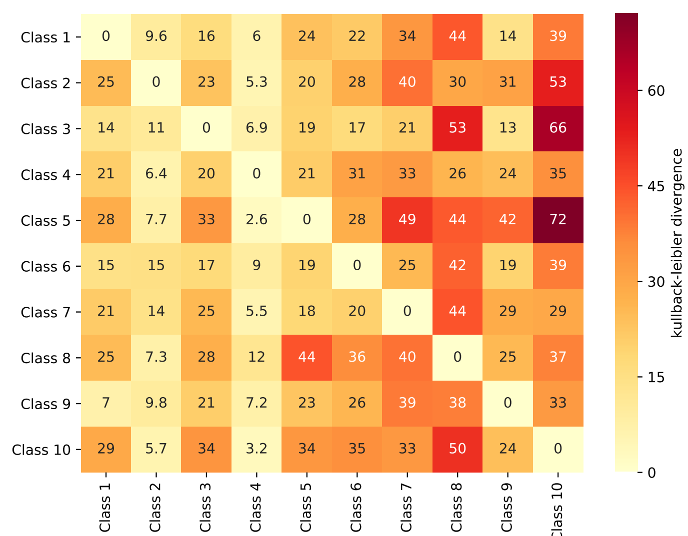


Figure 1. Distance found with KL-Divergence of Unigram for each class.

We first consider the distribution of linguistic features for each tourist spot. We compare Unigram, which is the probability distribution of words for each tourist spot. Tweets used for analyses were pre-processed using morphological analysis, for which we used Mecab [25]. Additionally, we used

 ISBN: 978-1-61208-690-3

TABLE I. TOP FIVE WORDS WITH HIGH BIAS IN UNIGRAM FOR EACH TOURIST SPOT.

Rank	Class 1	Class 2	Class 3	Class 4	Class 5
1	Taito Ward	Ota Ward	Shibuya Ward	Ota Ward	Ota Ward
2	Harajuku Station	Tokyo International Airport	Shibuya Station	Taito Ward	Tokyo International Airport
3	Tokyo International Airport	Passenger Terminal	Setagaya Ward	Tokyo International Airport	Passenger Terminal
4	Shibuya Ward	Chuo Ward	Tokyo International Airport	Passenger Terminal	Taito Ward
5	Tokyo Teleport Station	Haneda Airport	Tokyo Teleport Station	Tokyo Teleport Station	Haneda Airport
Rank	Class 6	Class 7	Class 8	Class 9	Class 10
1	Harajuku Station	Ota Ward	Tokyo Teleport Station	Ota Ward	Toyoshima Ward
2	Shibuya Ward	Taito Ward	Taito Ward	Nippon Budokan	Shibuya Ward
3	Jinbouchou Station	Minato Ward	Nippon Budokan	Minato Ward	Sunshine City
4	Taito Ward	Shinjuku Ward	Shibuya Ward	Sumida Ward	Nippori Station
5	Tokyo International Airport	Nippon Budokan	Tokyo International Airport	Tokyo Teleport Station	Arakawa Ward

the Good–Turing smoothed unigram language model in Kylm [26] for extracting Unigrams.

To examine the usefulness for classification with linguistic features, we divided tourist spots into 10 classes of $\{Class1, Class2, \cdots, Class10\}$ with class width of 10 based on the rank. In this study, our focus is not to reproduce the ranking of tourist spots on TripAdvisor completely, but to generate the ranking incorporating changes in popularity of tourist spots. Therefore, we set the number of classes to 10 for analyzing features that affect large ranking fluctuations rather than small fluctuations. Figure 1 is the result of comparing the difference of Kullback-Leibler Divergence (KL-Div) between Unigrams of linguistic features in each class. KL-Div has no symmetry which is an axiom of distance. Therefore, it can not be defined precisely as a distance. We confirmed the difference of KL-Div for each class. We assumed that linguistic features are useful for classification to a certain degree. Moreover, a large difference from $Class7$ to $Class10$ exists on average, which was regarded as caused by a special word distribution. Although we also compared KL-Div in each class by month, a similar tendency was apparent.

Table I presents results of comparing contribution rates of wordwise KL-Div in each class; then it displays the top five words of high bias. Many place names, station names, and facility names are apparent throughout all classes. Additionally, we confirmed a similar tendency even when verified by each month. As a cause for these results, we considered that many tourist spots are located in some areas belonging to a particular class. After examining the relation between the geographical factor belonging to a particular area and the ranking of tourist spots, we should consider whether a need exists to eliminate these geographical words during pre-processing. For the results described above, a certain difference is apparent between the distributions of words on linguistic features. In addition, results show that linguistic features contribute to the classification of ranking tourist spots. However, Twitter has many noise features, which implies that preprocessing and feature selection are extremely important.

C. Discussion of statistical features

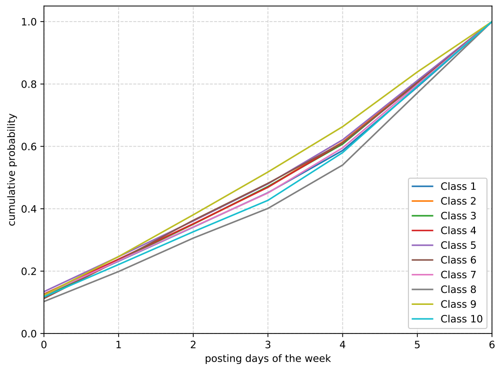
TABLE II. MEAN CORRELATION COEFFICIENT BETWEEN DIMENSIONS OF THE DISTRIBUTION OF STATISTICS FOR EACH TOURIST SPOT.

Date	Time	Character	Emoji	Emoticon	Post	Repeat	Revisit
0.52	**0.40**	0.22	0.29	**0.51**	0.20	0.22	**0.44**

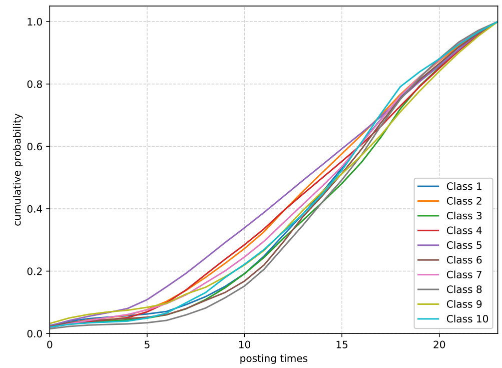
Next, we consider the distribution of statistical features for each tourist spot. We compare the cumulative probability distribution in the number of posting days of the week and times (hereinafter, Date and Time) and the number of posted characters, emoji, and emoticons (hereinafter Character, Emoji and Emoticon), which are statistics by tweets for each tourist spot. Similarly, we also compare the cumulative probability distribution in the number of posts, repeat posts and revisits (hereinafter Post, Repeat and Revisit), which are statistics by users on Twitter for each tourist spot.

To examine the usefulness for classification with statistical features, we divided tourist spots into 10 classes of $\{Class1, Class2, \cdots, Class10\}$ similarly to the process described in Section III-B. Figure 2 is the result of comparing the cumulative probability distribution of statistical features in each class. Figure 2a presents the cumulative probability distribution of each class. The horizontal axis shows days of the week, from Monday through Sunday. We confirmed distinct differences between the weekdays (0–4) and holidays (5–6) from data shown in Figure 2a. Figure 2b portrays the cumulative probability distribution of posting times of tweets posted in each class. The horizontal axis shows 24 hours in a day. A similar tendency is apparent from Figure 2a between the daytime (6–18) and night time (0–5, 18–23). Therefore, we consider that these features make a certain contribution to classification. Figure 2c displays the cumulative probability distribution of the number of posted characters in each class. The distribution in each class is complicated. Therefore, it remains unclear whether this feature contributes to classification. Figure 2d exhibits the cumulative probability distribution of the number of posted emoji in each class. Because of differences in each class, we can infer that this feature contributes to classification to a certain degree. However, a high bias is apparent by the number of occurrences of 0 throughout all classes. Figure 2e depicts the cumulative probability distribution of the number of posted emoticons in each class. We were unable to see differences in each class. Therefore, this feature is not regarded as contributing to classification.

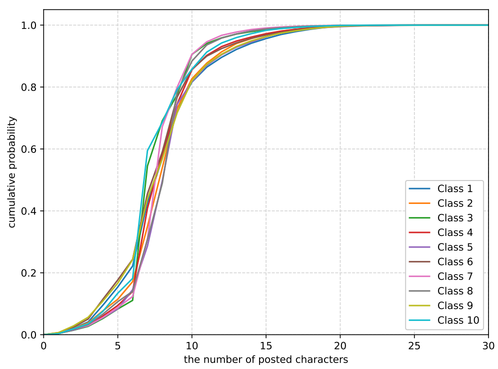
Figure 2f is the cumulative probability distribution of the number of posts by user on Twitter in each class. We were able to detect a certain difference in each class. However, a high bias exists by the number of occurrences of 1, similarly to Figure 2d. Figure 2g is the cumulative probability distribution of the number of repeat posts by users on Twitter in each class. There are differences in each class and a high bias to the number of occurrences of 0 similarly in Figure 2d and Figure 2f. Figure 2h is the cumulative probability distribution of the number of revisits by users on Twitter in each class, which closely resembles Figure 2g. From the above, we were able to confirm several differences in the distributions of



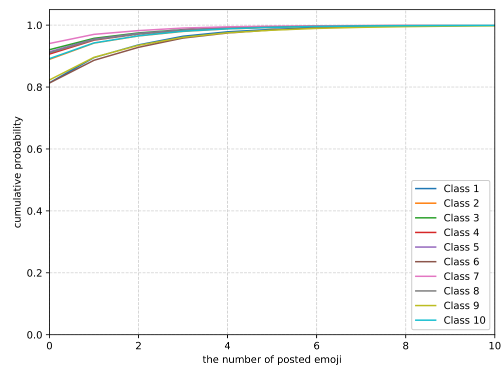
a. Cumulative distribution functions based on posting days of the week of tweets.



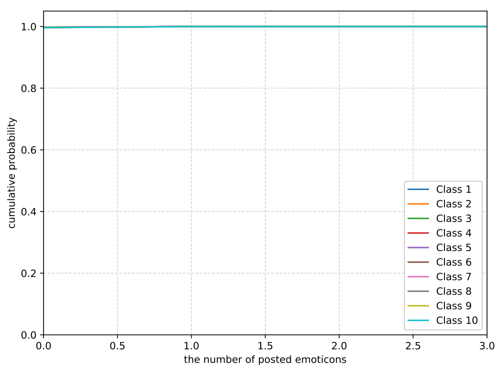
b. Cumulative distribution functions based on posting times of tweets.



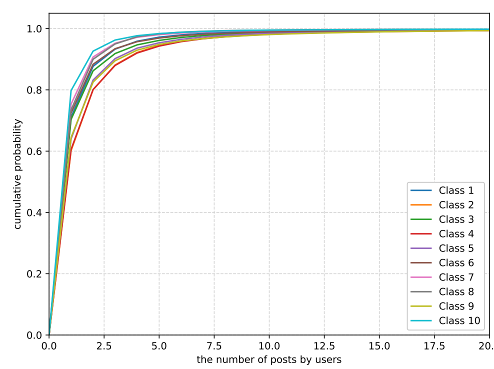
c. Cumulative distribution functions based on the number of posted characters on tweets.



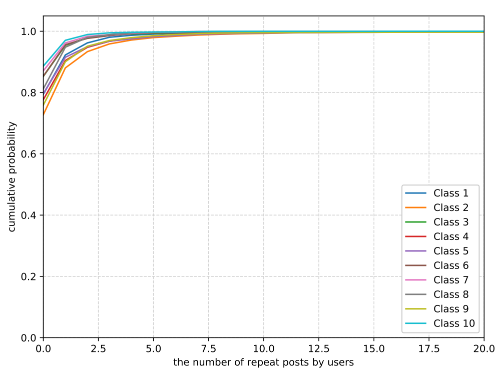
d. Cumulative distribution functions based on the number of posted emoji on tweets.



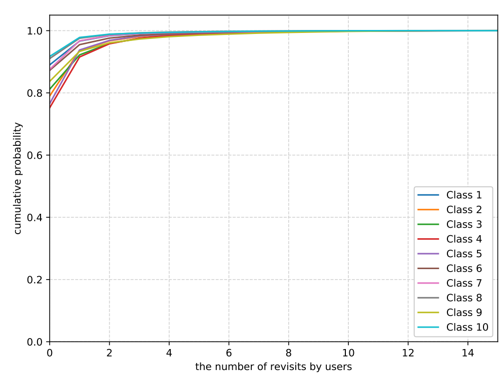
e. Cumulative distribution functions based on the number of posted emoticons on tweets.



f. Cumulative distribution functions based on the number of posts by users on Twitter.



g. Cumulative distribution functions based on the number of repeat posts by users on Twitter.



h. Cumulative distribution functions based on the number of revisits by users on Twitter.

Figure 2. Cumulative distribution functions of the statistical features for each tourist spot.

statistics by tweets and Twitter users. However, a high bias exists in the distribution of statistics by Twitter users. We must consider the influence of classification.

Table II shows mean correlation coefficients between dimensions of the features in each statistic. The criterion of correlation is 0.4, which is generally regarded as representing some degree of correlation. We should consider the possibility of negative influence on classification by multicollinearity because of a high mean correlation coefficient between dimensions in Date, Time, and Revisit. As described earlier, a certain difference exists between the distribution of the statistical features. We consider that several statistics, such as Date and Time, will make some contribution to classification. However, we must devote some attention to the influence of multicollinearity in classification.

IV. Experiment

In this section, to model the ranking of tourist spots on TripAdvisor for generating the ranking using tweets, we used 50 tourist spots in Section III-A. First, we analyze features by logistic regression and perform ranking learning using the result.

A. Classification

TABLE III. COMPARISON OF SUBSET ACCURACY BY LOGISTIC REGRESSION WITH LINGUISTIC FEATURES.

Features	Accuracy (+/- Error Rate)
Unigram	0.90 (+/- 0.02)
Unigram + IG	0.71 (+/- 0.02)
Unigram + TFIDF	**0.92 (+/- 0.01)**
Bigram	**0.94 (+/- 0.01)**
POS Unigram	0.26 (+/- 0.03)
Word2Vec	0.38 (+/- 0.03)

TABLE IV. TOP 20 WORDS WITH HIGH CONTRIBUTION RATES OF SPEARMAN'S RANK CORRELATION COEFFICIENT (SRCC) OF 0.4 OR MORE IN CONTRIBUTION RATES BY LOGISTIC REGRESSION.

Word	Coefficient
Collection	**0.6261**
Beautiful	**0.6071**
Walk	0.5860
Doing	0.5784
Fall	0.5599
Amazing	**0.5342**
Meal	0.5227
Buy	0.4644
Last night	0.4512
Meet	0.4497
Customer	0.4369
Good	**0.3802**
Many	0.3717
Flow	0.3438
Combination	0.3294
Ballet	0.3273
Old man	0.3082
Workplace	0.2948
Near	0.2648
Sit	0.2357

Here, we perform logistic regression and analyze effective features for ranking learning from contribution rates. We used the scikit-learn [27] algorithm for the implementation of logistic regression. To generate the highest performance model, we performed a grid search with five cross validation for hyperparameters.

TABLE V. COMPARISON OF SUBSET ACCURACY BY LOGISTIC REGRESSION WITH STATISTICAL FEATURES.

Features	Accuracy (+/- Error Rate)
All Statistics	0.39 (+/- 0.03)
All Statistics - Emoji - Emoticon	0.40 (+/- 0.03)
Tweet Statistics	0.58 (+/- 0.02)
User Statistics	0.25 (+/- 0.04)
Date + Time + Character	**0.64 (+/- 0.05)**

We used linguistic features and statistical features for logistic regression. For this experiment, we used four linguistic features of Unigram, Bigram, Part-Of-Speech (POS) Unigram, and Word2Vec. However, we performed feature selection with Information Gain (IG) and Term Frequency-Inverse Document Frequency (TFIDF) to Unigram by reducing dimensions of the features below each mean value. As statistical features, we used statistics by tweet (hereinafter, Tweet Statistics), statistics by users on Twitter (hereinafter, User Statistics) and All Statistics combining them in Section III-A.

For classification with logistic regression, we divided tourist spots into 10 classes of $\{Class1,\ Class2,\ \cdots,\ Class10\}$ with the class width of 10 based on the rank. Table III presents the subset accuracy of logistic regression with linguistic features. This subset accuracy takes 1 when the set of labels predicted by the test data exactly matches the set of labels on the answer data in the multi-class classification. As a result, the model generated by Unigram + TFIDF and Bigram has the highest performances in linguistic features. Because 3,521 dimensions of Unigram + TFIDF are far fewer than the 1,203,938 dimensions of Bigram, we infer that the effective features on Unigram + TFIDF were extracted by feature selection for classification.

Table IV shows the top 20 words of classification contribution rates for which Spearman's Rank Correlation Coefficient (hereinafter, SRCC) is 0.4 or more in the classification model in Unigram + TFIDF. As with Section III-B, we set the criterion of correlation to 0.4. We were able to confirm that "Collection", which seems to express a short-term event and "Beautiful" and "Good", which are favorable impressions of tourist spots, are placed higher in the ranking. In the case of comparing words solely by contribution rates, we confirmed that many geographical words appear in the top of the ranking. We consider that feature selection by rank correlation coefficient is useful to a certain degree.

Table V presents the accuracy of logistic regression with statistical features. As described in Section III-B, some features, such as Emoji and Emoticon, do not contribute to the classification of ranking of tourist spots. In classification with tweet statistics and user statistics, we confirm that the classification with tweet statistics shows high performance. Moreover, the higher accuracy of classification was recorded by eliminating Emoji and Emoticon based on Section III-B. As a result, Date + Time + Character, which differed in each class and which were free from bias in Section III-B exhibited the highest performance. As described previously, we consider that user statistics did not make a contribution to classification because of high bias. Next, we will model the ranking of tourist spots by ranking learning after selecting features with a high rank correlation coefficient of classification.

 ISBN: 978-1-61208-690-3

TABLE VI. COMPARISON OF PAIRWISE ACCURACY AND MEAN NDCG BY THE RANKING LEARNING MODEL WITH FEATURES: LF, UNIGRAM + TFIDF; SF, DATE + TIME + CHARACTER.

Features	Pairwise Accuracy	Mean NDCG
LF	0.8864	0.8918
SF	0.6666	0.1741
LF + SF	0.8329	0.7692
LF (SRCC $>$ 0.4)	0.7915	**0.8998**
SF (SRCC $>$ 0.4)	0.6440	**0.2038**
LF + SF (SRCC $>$ 0.4)	0.8528	**0.8075**

B. Ranking Learning

From the result presented above, we performed ranking learning with linguistic features and statistical features to model the ranking of tourist spots on TripAdvisor. We used RankSVM, which is a ranking learning algorithm incorporating Kendall's rank correlation coefficient to Support Vector Machine, for which we adopted the algorithm of LIBSVM [28] proposed by Lee et al. [29]. To generate the highest performance model, we also performed a grid search with five cross validation for hyperparameters.

As features for RankSVM, we used effective features for classification in Section IV-A, which is Unigram + TFIDF in Linguistic Features (hereinafter, LF) and Date + Time + Character in Statistical Features (hereinafter, SF). Additionally, we verify the usefulness of SRCC for ranking learning in each feature.

Table VI shows pairwise accuracy and mean Normalized Discounted Cumulative Gain (hereinafter, NDCG) of RankSVM in each features. NDCG takes a value from 0 to 1. It is closer to 1 when the result of ranking conforms to answer data, which is a widely used prediction result indicator of ranking method. Additionally, we apply selected features for which SRCC is higher than 0.4 in Table IV. Results show that the values of NDCG were improved by feature selection with SRCC in both linguistic features and statistical features. We also confirmed a high performance of ranking learning while combining both features. As a result, the model generated by ranking learning with LF exhibited the highest performance.

V. Discussion

We performed ranking learning by considering SRCC of contribution rates by logistic regression. In terms of linguistic features, from the accuracy of classification in Table III, large differences exist among tourist spots. Many geographical words were ranked higher when comparing the features solely by contribution rates according to Section III-A. However, considering feature selection by rank correlation coefficients, we were able to confirm that words which give positive impressions to tourist spots were ranked higher in Table IV.

We also discussed statistical features by tweets and Twitter users. Regarding contribution rates, it is apparent that Date, Time, and Character, which have statistics by tweets, have a high contribution rate. However, we confirmed certain variations in the distribution of Post, Repeat and Revisit in Section III-B. We consider that these statistics did not contribute to classification because the bias to the number of occurrences of 0 or 1 is too large. Additionally, we confirmed a certain correlation for dimensions of the distribution of Date and Time in Table II. Regarding multicollinearity, this is regarded as negatively influencing the accuracy of classification. Therefore, if we were to have sparse aggregation of statistics, then we might improve the accuracy of classification with statistical features.

In ranking learning, we performed feature selection based on SRCC of contribution rates by logistic regression. Consequently, NDCG recorded high performance for both linguistic features and statistical features. We infer that feature selection by a rank correlation coefficient of contribution rates is useful to a certain degree. Both pairwise accuracy and mean NDCG were also the highest scores with linguistic features. As described previously, the case of performing precise preprocessing demonstrates that linguistic features are extremely important for modeling the ranking of tourist spots.

Table VII presents the result of applying recent tweets to the model in ranking learning by LF and LF (SRCC $>$ 0.4), which are high performance in Section VI. We confirm that more tourist spots hold seasonal events at the top of the ranking by Twitter than by TripAdvisor. For instance, "The National Art Center, Tokyo", "Sunshine City" and "Tokyo International Forum" are in the top three of the ranking. These spots hold exhibitions, events or other activities frequently. Additionally, there are some differences in ranking from the viewpoint of rank correlation for feature selection. When the contents of these increase, we can not determine if it is profitable. Therefore, it is necessary to conduct an experiment by users.

As a result, we could confirm that linguistic features are more effective for ranking of tourist spots incorporating changes in popularity than statistic features. We speculate that this is because the number of dimensions of our statistic features is not enough to deal with this problem. Finally, the result of our experiment demonstrated that tourist spots which hold events frequently are ranked higher by applying recent tweets to the model in ranking learning. Although further experiments would be required, we could also confirm that the values of NDCG are improved by considering SRCC of contribution rates.

VI. Conclusion and Future Work

As described in this paper, we generated a ranking that incorporates popularity variation by modeling the ranking of tourist spots on TripAdvisor using tweets. Twitter has numerous noise posts. Therefore, we performed feature selection by a rank correlation coefficient of contribution rates by logistic regression. Results demonstrate that we were able to improve the performance of the model in ranking learning. Eventually, the model by RankSVM showed the highest NDCG of 0.89. However, pairwise accuracy decreased from the highest score of 0.88 to 0.79.

Future work must address the contribution rates themselves by logistic regression for feature selection. As described in this paper, we used only rank correlation coefficients for feature selection because the contribution rate of geographical words to classification is too large. Therefore, we expect improvement of classification accuracy in ranking learning by eliminating geographical words during pre-processing. Additionally, it is necessary to consider evaluation after proposing ranking by the proposed method to the user. From results of demonstration experimentation, we expect to verify whether tourist spots to enjoy most at the time of visit are suggested in the ranking, which is information that users actually need.

TABLE VII. TOP TEN TOURIST SPOTS BY THE RANKING OF TRIPADVISOR AND TWITTER.

Rank	TripAdvisor	Twitter	Twitter (SRCC $>$ 0.4)
1	Ryogoku Kokugikan	The National Art Center, Tokyo	Sunshine City
2	Asakusa	Sunshine City	The National Art Center, Tokyo
3	Roppongi Hills	Tokyo International Forum	Tokyo International Forum
4	Chidorigafuchi	Bunkamura	Bunkamura
5	Happoen Garden	Roppongi	Ameya Streets
6	Meiji Jingu Shrine	Yanaka Cemetery	Spa LaQua
7	Tokyo Camii & Turkish Culture Center	Odaiba Palette Town	Omoide Yokocho
8	Senso-ji Temple	Ginza Namiki Streets	Roppongi
9	Tokyo Metropolitan Government Buildings	Ryogoku	Tokyo Character Streets
10	Tokyo National Museum	Kitanomaru Park	Imperial Palace

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Numbers 16K00157 and 16K16158, and a Tokyo Metropolitan University Grant-in-Aid for Research on Priority Areas "Research on Social Big Data."

REFERENCES

[1] TripAdvisor, "TripAdvisor: Read Reviews, Compare Prices & Book." [Online]. Available: https://www.tripadvisor.com/ [Accessed: Jan., 2019].

[2] Z. Xiang, D. Wang, J. T. O' Leary, and D. R. Fesenmaier, "Adapting to the internet: trends in travelers' use of the web for trip planning," Journal of Travel Research, vol. 54, no. 4, 2015, pp. 511–527.

[3] U. Gretzel and K. H. Yoo, "Use and impact of online travel reviews," Information and communication technologies in tourism 2008, 2008, pp. 35–46.

[4] Twitter, "Twitter. It's what's happening." [Online]. Available: https://twitter.com/ [Accessed: Jan., 2019].

[5] Facebook, "Facebook." [Online]. Available: https://www.facebook.com/ [Accessed: Jan., 2019].

[6] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," in Proceedings of the 20th international conference on World wide web. ACM, 2011, pp. 675–684.

[7] R. Filieri, S. Alguezaui, and F. McLeay, "Why do travelers trust tripadvisor? antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth," Tourism Management, vol. 51, 2015, pp. 174–185.

[8] A. Mukherjee, V. Venkataraman, B. Liu, and N. S. Glance, "What yelp fake review filter might be doing?" in ICWSM, 2013, pp. 409–418.

[9] I. Guy, A. Mejer, A. Nus, and F. Raiber, "Extracting and ranking travel tips from user-generated reviews," in Proceedings of the 26th international conference on world wide web. International World Wide Web Conferences Steering Committee, 2017, pp. 987–996.

[10] J. Borràs, A. Moreno, and A. Valls, "Intelligent tourism recommender systems: A survey," Expert Systems with Applications, vol. 41, no. 16, 2014, pp. 7370–7389.

[11] M. Ye, P. Yin, and W.-C. Lee, "Location recommendation for location-based social networks," in Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems. ACM, 2010, pp. 458–461.

[12] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee, "Exploiting geographical influence for collaborative point-of-interest recommendation," in Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval. ACM, 2011, pp. 325–334.

[13] Foursquare, "Foursquare." [Online]. Available: https://foursquare.com/ [Accessed: Jan., 2019].

[14] D. Choudhury et al., "Automatic construction of travel itineraries using social breadcrumbs," in Proceedings of the 21st ACM conference on Hypertext and hypermedia. ACM, 2010, pp. 35–44.

[15] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, "Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency," Knowledge and Information Systems, vol. 54, no. 2, 2018, pp. 375–406.

[16] Flickr, "Find your inspiration. - Flickr." [Online]. Available: https://www.flickr.com/ [Accessed: Jan., 2019].

[17] S. Ishihara, M. Nagamachi, and T. Tsuchiya, "Development of a kansei engineering artificial intelligence sightseeing application," in International Conference on Applied Human Factors and Ergonomics. Springer, 2018, pp. 312–322.

[18] Y. Mizutani and K. Yamamoto, "A sightseeing spot recommendation system that takes into account the change in circumstances of users," ISPRS International Journal of Geo-Information, vol. 6, no. 10, 2017, p. 303.

[19] S. Missaoui, M. Viviani, R. Faiz, and G. Pasi, "A language modeling approach for the recommendation of tourism-related services," in Proceedings of the Symposium on Applied Computing. ACM, 2017, pp. 1697–1700.

[20] Y. Duan, L. Jiang, T. Qin, M. Zhou, and H.-Y. Shum, "An empirical study on learning to rank of tweets," in Proceedings of the 23rd International Conference on Computational Linguistics. Association for Computational Linguistics, 2010, pp. 295–303.

[21] A. Gupta and P. Kumaraguru, "Credibility ranking of tweets during high impact events," in Proceedings of the 1st workshop on privacy and security in online social media. ACM, 2012, p. 2.

[22] Y. Chang et al., "Improving recency ranking using twitter data," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 4, no. 1, 2013, p. 4.

[23] Google, "Places - Google Maps Platform - Google Cloud." [Online]. Available: https://cloud.google.com/maps-platform/places/ [Accessed: Jan., 2019].

[24] OpenStreetMap, "OpenStreetMap Nominatim: Search." [Online]. Available: https://nominatim.openstreetmap.org/ [Accessed: Jan., 2019].

[25] Kyoto University, "MeCab: Yet Another Part-of-Speech and Morphological Analyzer." [Online]. Available: http://taku910.github.io/mecab/ [Accessed: Jan., 2019].

[26] G. Neubig and X. Yao, "Kylm - Kyoto Language Modeling Toolkit." [Online]. Available: http://www.phontron.com/kylm/ [Accessed: Jan., 2019].

[27] scikit-learn, "scikit-learn: machine learning in Python." [Online]. Available: http://scikit-learn.org/stable/ [Accessed: Jan., 2019].

[28] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," ACM Transactions on Intelligent Systems and Technology, vol. 2, 2011, pp. 27:1–27:27.

[29] C.-P. Lee and C.-J. Lin, "Large-scale linear ranksvm," Neural computation, vol. 26, no. 4, 2014, pp. 781–817.