FSDM 2020
November 13-16, 2020
Online Conference

FSDM3409

Keyword Extraction from TV Program Viewers' Tweet based on Neural Embedding Model

Taiga KIRIHARA[a] , Kazuyuki MATSUMOTO[b] , Minoru YOSHIDA[b], Kenji KITA[b]

[a] Graduate Schools of Science and Technology for Innovation Division of Science and Technology, Tokushima University, Japan

[b] Graduate School of Technology, Industrial and Social Sciences, Tokushima University, Japan

1. Introduction

Problem

The number of people who watch television (TV) programs by recording them has increased in recent years.

↓

・The program is long
・Limited to viewing time

↓

Difficult to watch the entire program

The solution shown in this study

More viewers tweet while watching TV programs on Twitter.

↓

Estimate the event scene from tweets about the program content tweeted during the program broadcast.

↓

Reduce viewing time

2. Proposed Method

Tweet collection and vectorization

- Collect the tweet text containing "# [program name]" using the Twitter API
- Divide the tweet set of each program into 1 minute units
- Vectorization using BERT, is trained model on the Japanese Wikipedia corpus

Input: *Raishu no Anipoke wa Kabigon kyodaika desu !*

↓ Morphological analysis with MeCab+Neologd

Raishu / no / Anipoke / wa / Kabigon / kyodai / ka / desu / !

With the part-of-speech pattern of "proper noun + adjective" as the condition, phrases that match it are candidates

↓

Vector is calculated by calculating sentence embedding for candidate phrases and entire sentences

↓

Candidate phrases close to the vector of the document itself are extracted based on the degree of similarity, and key phrases are determined.

↓

Output: *Anipoke, Kabigon*

BERT

- BERT is a deep learning model for natural language processing that by Google in 2018.
- Compared to Word2vec, it has the feature that the entire sentence can be vectorized.

Key phrase extraction

Many combinations of "noun + adjective" can be seen in live tweets.

↓

- The EmbedRank algorithm mainly extracts patterns of "proper nouns" + "adjectives"
- The EmbedRank algorithm extracts candidate phrases extracted based on part of speech patterns and important phrases based on sentence vectors.

Key phrase extraction example

- "The giant Snorlax will appear in the next week's Anipoke!" (Raishu no Anipoke wa Kabigon kyodaika desu.)

↓

"Anipoke" and "Snorlax"

- "Pikachu is cute today too !" (Kyou mo Pikachu kawaii !)

↓

"Pikachu" and "cute"

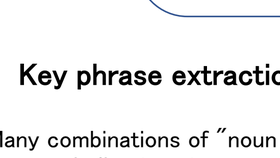

Figure 1. Pikachu in "Pokémon"

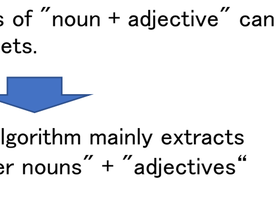

Figure 2. Pennywise in "It"

MMR method

The MMR method is used because similar expressions are duplicated by simply extracting the phrase.

$$MMR := \arg\max_{D_i \in R \setminus S} \left[\lambda \cdot Sim_1(D_i, Q) - (1-\lambda) \max_{D_j \in S} Sim_2(D_i, D_j) \right]$$

Scene classification

Manually classify scenes into 6 categories.

- **People**: Scenes in which people appear
- **Action**: A scene in which a person takes action
- **Emotion**: A scene in which a person shows emotion
- **Summary**: Rough description of a scene without humans
- **Scenery**: A scene where the location and time of day can be observed
- **Conversion**: A change of scene from the main program to other images (e.g., a commercial)

3. Data

- Collect tweet data of 3 programs of "Variety", "Movie" and "Anime".
- The collection time is from 10 minutes before the start of the program to 10 minutes after the end of the program.
- Prepared the correct answer data attached manually.

Table 1. Details of the collected programs

Broadcaster	On-air date	Program name	Number of tweets
Nippon Television	2019/10/14	Getsuyou Kara Yofukashi	14239
Nippon Television	2019/11/8	Friday Road SHOW – It	19460
TV TOKYO	2019/12/8	Pokémon	4105

Table 2. Number of scenes in each program category

Category	Yofukashi	Friday	Pokémon	Total
Person	5	13	7	25
Action	8	70	4	82
Emotion	1	1	1	3
Summary	7	2	0	9
Scenery	2	19	2	23
Conversion	26	17	9	52
Total	49	122	23	194

4. Result

Experimental method:

- Using the correct answer data as a search query, calculate the top 5 time zones based on the similarity with the key phrase.
- t-SNE is used to compress the dimensions and visualize them on the two-dimensional coordinate space.

Evaluation method:

- Compare the time of correct data and the error of output result
  - Calculate difference between the correct time and the predicted time
- Also, compare with MMR and without MMR

Experimental result:

Event scene search results:
Less error with MMR than without MMR.

Table 3. Average error of each program

	With MMR	Without MMR
Yofukashi	9 min. 41 sec.	14 min. 43 sec.
Friday	11 min. 3 sec.	16 min. 3 sec.
Pokémon	3 min. 33 sec.	5 min. 38 sec.

Key phrase visualization for each scene category in "Friday Road SHOW":

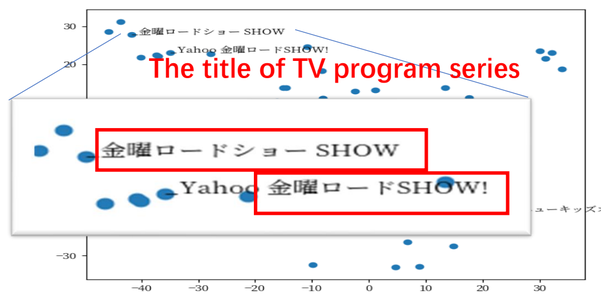

Figure 3. Key phrases in the "person" scene

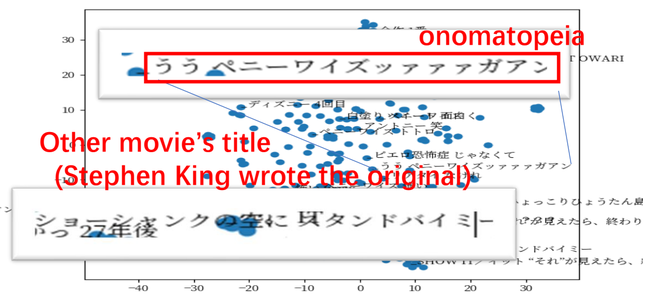

Figure 4. Key phrases in the "action" scene

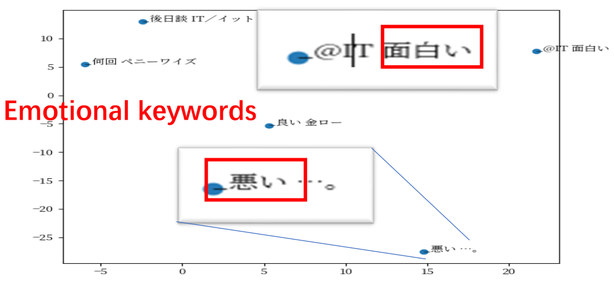

Figure 5. Key phrases in the "emotion" scene

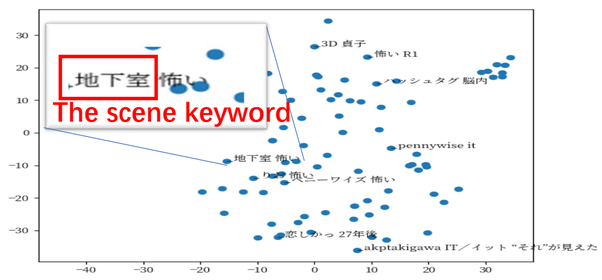

Figure 6. Key phrases in the "scenery" scene

5. Conclusion

Event scene search

- about 5% to 10% of the total error that was observed in the performance evaluation was due to the time zone error.
- MMR was able to suppress duplication of similar key phrases.

- Difference between correct answer data and tweet text.
- We used a Wikipedia-based trained model, but We assume We couldn't handle internet slang such as "Kita" well.

Scene classification and visualization

- Since many **proper nouns** and **adjectives** are included in the tweet, the result is that many emotional words and nouns are included in the "**action**" scene classification.
- Even in the "**emotion**" and "**summary**" scenes, the verbs, adverbs, and general nouns that are necessary to describe the scene could not be extracted.

The next deployment

- Use pre-learned BERT based on SNS
- Weighting key phrases
- Expansion of key phrase candidates by optimizing part of speech conditions

*This work was supported by JSPS KAKENHI Grant Numbers JP20K12027, JP18K11549.