Classification of Nostalgic Music Through LDA Topic Modeling and Sentiment Analysis of YouTube Comments in Japanese Songs

Kongmeng Liew[1], **Yukiko Uchida**[2], **Nao Maeura**[1], and **Eiji Aramaki**[3]

[1]Graduate School of Human and Environmental Studies, Kyoto University, Japan
[2]Kokoro Research Center, Kyoto University, Japan
[3]Graduate School of Science and Technology, Nara Institute of Science and Technology, Japan
[1]{ryuu.konmin.56x, maeura.nao.35x}@st.kyoto-u.ac.jp
[2]uchida.yukiko.6m@kyoto-u.ac.jp
[3]aramaki@is.naist.jp

Abstract

Nostalgia has been defined as a bittersweet, social emotion, that is often induced through music. In this paper, we examine how these may be expressed in Japanese YouTube comments of nostalgic (mid-2000s) and non-nostalgic (recent) songs (music videos). Specifically, we used sentiment analysis and Latent Dirichlet Allocation (LDA) topic modeling to examine emotion word usage and broader themes across comments. A gradient boosted decision tree classifier was then able to classify nostalgic and non-nostalgic music videos above chance level. This suggests that analyses on video/music comments may be a possible method to quantify expressions of listener emotions, and categorise musical stimuli.

1 Introduction

The last decade has seen a sharp increase of nostalgia-related research in the psychology-emotion literature. Nostalgia has been defined primarily as a self-relevant emotion, in that the self is experienced through narratives of autobiographical events. Yet, it is also a social emotion, in that these narratives also involve memories of social interaction, ultimately fostering a sense of social connectedness (Tilburg et al., 2017, 2018; Reid et al., 2015; Vess et al., 2012). It has been characterized as a bittersweet experience, mixing feelings of pleasantness with appraisals of irretrievable loss (Tilburg et al., 2018), particularly in reflecting and savouring past social experiences (Biskas et al., 2019). In music, it is often induced by sadness (Taruffi and Koelsch, 2014), and is stronger for music associated with reminiscence bumps (i.e., disproportionally recalled for events in late adolescence and early childhood, Krumhansl and Zupnick, 2013)

In this paper, we propose that since nostalgia has such distinct elicitors and appraisals, automatic classification of nostalgic popular songs should be possible by analysing listener responses. Here, we operationalise these responses as comments on music videos in YouTube. We first use unsupervised learning (topic modelling) and sentiment analysis to quantify comments into features, and use supervised learning (gradient boosted decision trees (GBDT, Friedman, 2001)) to classify comments belonging to nostalgic (old) music videos, or non-nostalgic (recent) music videos based on the identified topics and sentiment categories.

2 Related Work

In the field of Music Information Retrieval, social media (Twitter posts) has been previously used in the context of music entity recognition. Porcaro and Saggion (2019) developed a method of identifying aspects of broadcast classical music through corresponding Twitter activity. For nostalgia and music, Timoney, Davis and Raj (2018) mined 556 comments from YouTube music videos from British hit songs between 1960 – 1970, and found that nostalgic comments could be classified with 86% accuracy (from non-nostalgic comments). This mirrors research from Davalos and colleagues (2015), who found distinctive characteristics of nostalgic posts on Facebook: nostalgic posts tended to have more reflective and emotional content, tinged with mixed positive and negative elements. Our analysis adds to this body of research, in that we seek to use classify comments belonging nostalgic and non-nostalgic music videos in Japanese.

3 Method

We first conducted an online pilot study, where N(participants) = 342 participants rated one randomly selected song (out of a total set of 20 songs)

Track	Artist	Condition	Year
No More Cry	D-51	Nostalgic	2005
Kibun Jou Jou	Mihimaru GT	Nostalgic	2006
Goodbye Days	YUI	Nostalgic	2007
Sakura	Naotaro Moriyama	Nostalgic	2002
Wataridori	[Alexandros]	Non-Nostalgic	2015
Stay Tune	Suchmos	Non-Nostalgic	2017
Chocho Musubi	Aimer	Non-Nostalgic	2016
Himawari no Yakusoku	Motohiro Hata	Non-Nostalgic	2015

Table 1: List of songs per condition and release year. YouTube IDs for each video are available in our online supplementary material.

on felt nostalgia. Each song received ratings from approximately 10 participants. From this, we selected 8 songs that scored the highest (and lowest) on felt nostalgia via a single-item, 7-point Likert scale. We defined nostalgic songs as songs that were popular within Japan in the mid-2000s, that likely induced nostalgia for those aged around 25-35. Non-nostalgic songs were recently popular songs released within the last 5 years (see Table 1). We then identified 37 YouTube videos that corresponded to these 8 songs and obtained a list of all YouTube comments through the YouTube API via the 'tubeR' wrapper in R (Sood, 2019). To ensure the overall representativeness of our study, this excluded videos that had less than 50000 views, were not in Japanese, and collaboration videos. Additionally, we filtered out exceptionally short comments (that were deemed unsuitable for analysis), by excluding the shortest (25th percentile) comments from the dataset. We also removed all alphanumeric characters and non-Japanese text, and tokenised the remaining Japanese comments through the RMeCab (Ishida, 2018) wrapper for the MeCab software (Kudo, 2005). This converted Japanese terms and phrases into their simplest (plain) forms, allowing for more consistency in both topic modelling, and matching with the emotion dictionary. We obtained a final N(comments) = 710 (Nostalgic = 324, Non-Nostalgic = 386). We obtain scores for emotion tags and LDA posterior probabilities for all comments, and divided them into training (0.6) and testing (0.4) sets. We used GBDT ('gbm' package; Greenwell et al., 2019, using the 'caret' wrapper; Kuhn, 2019) to classify them as nostalgic or non-nostalgic, and use partial dependency plots and variable importance measures to interpret these results. Note that all analyses were conducted in R (R Core Team, 2019).

3.1 Text Analyses

For the sentiment analyses, we used the JIWC emotion dictionary (Shibata et al., 2017). This matched words to 7 emotion categories (happy, sad, anger, surprise, trust, anxiety, hate/disgust; fear was excluded) based on a translation of Pluchik's (1980) emotion wheel, and scores for each comment (S_{ij}) were a ratio of number of emotion terms in each category (W_{ij}), to the total number of terms (tokens; W_i*) in each comment:

$$S_{ij} = \frac{W_{ij}}{W_i*} log(W_{ij} + 1) \qquad (1)$$

For topic modelling, in order to reduce the bias caused by human supervision, this study employed the unsupervised Latent Dirichlet Allocation (LDA; Blei et al., 2003). LDA identifies latent topics from documents (in this case, comments), through modelling the probabilistic distribution of topics in a document, and words in topics. LDA topic modelling with Gibbs sampling was conducted used the 'topicmodels' package (Grun and Hornik, 2011), and the number of topics was determined used the method described in Griffiths and Steyver (2004), which uses the posterior probability for each model (with varying numbers of topics) from all words in the corpus of YouTube comments. For each document, we used the resultant probability distribution for each topic as features in a classification model alongside the JIWC emotion categories.

4 Results

A total of 14 topics were identified (see Figure 1; a list of top terms for all topics are available in our online supplementary material: https://osf.io/52abe/). These were combined with scores from the 7 JIWC emotion categories and word count, and fitted in a GBDT classification model, for a total of 22 features. Parameter selection for the model was determined through 12-fold cross validation on the training set, resulting in n(trees) = 50 and interaction depth = 1. An overall modest accuracy score of AUC = 0.60, Mc Neymar $p < .001$ was achieved when fitted on the test set. This suggested that the model was weakly but significantly able to classify nostalgic and non-nostalgic songs based on YouTube comments above chance-level. As such, we believe that small but significant differences exist between comments from nostalgic and non-nostalgic songs.

To understand what these features were and how they affected classification in nostalgic and non-nostalgic songs, we interpreted the model through permutation feature importance (PFI), and partial dependence plots (PDP)s by the 'iml' (Molnar et al., 2018) and 'pdp' (Greenwell, 2017) packages (all PFI scores are available in our online supplementary material: https://osf.io/52abe/). We instituted a cutoff of importance = 1.01 for PFI, which selected 5 features of importance for interpretation. These were Topics 4, 1, 14, and 13, as well as the JIWC-Happy emotion category. The PDPs revealed that Topics 4, 1, 14, and Happy were higher in nostalgic music comments, but Topic 13 displayed an inverted-U relationship.

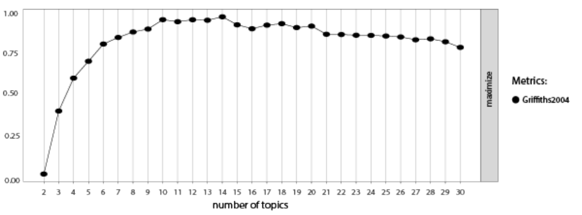

Figure 1: LDA Model likelihood at different numbers of topics for model selection.

5 Discussion

We labelled Topic 4 as 'Bittersweet', as it contained words that expressed both happiness and sadness, that appear self-directed and focused (e.g., 'happiness', 'tears', 'self', 'believe', 'can-do', and 'find'). Topic 1 included several self-directed, high-arousal words, such as 'live (music)', 'the best', 'favourite', and 'cool', that we labelled as 'High-arousal'. Topic 13 consisted of several words like 'courage', 'sitting for entrance exams', and 'striving', so we labelled it as 'Entrance Exams', and Topic 14 included words like 'good', 'family', 'children', that we labelled

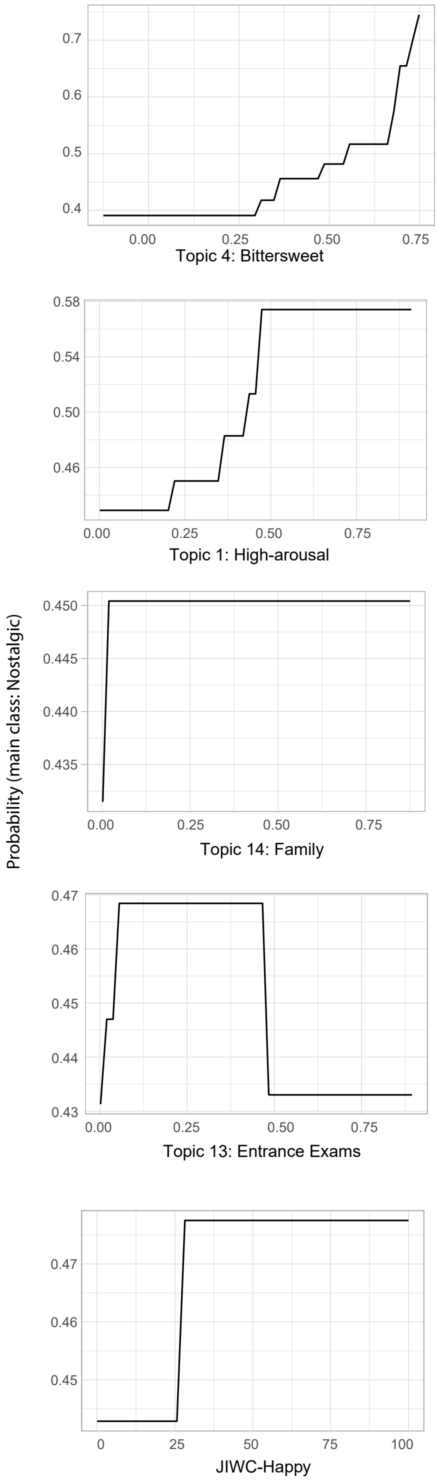

Figure 2: PDPs for high importance features (topics): the larger the probability (y-axis), the higher the probability of classification as nostalgic music. X-axis indicates the posterior probability for each topic or JIWC emotion category frequency scores

as 'Family'. These topics, as well as happiness-related words, had an influence on the model in classifying comments. However, Topic 13 'Entrance Exams' displayed a somewhat inconsistent relationship, in that the comments of low to mid probabilities on that topic were more likely to belong to nostalgic songs, but comments which were very low, and also high on that topic were from non-nostalgic songs.

Nevertheless, we conclude that comments on nostalgic songs had a greater likelihood of mentioning topics that related to bittersweet and/or high-arousal emotions, and happiness. Furthermore, they included mentions of social memories (such as family), and to a certain extent, collective memory (such as sitting for entrance exams - commonly considered a rite of passage in Japanese youth). These appear to be consistent with previously-identified appraisals and construals of Nostalgia in past literature (Sedikides and Wildschut, 2019; Tilburg et al., 2018)

However, we note the low classification accuracy of the model. It is likely that newer, more powerful models, like Latent Feature topic modeling (LFTM) and Long Short-Term Memory (LSTM) neural network classifiers, and larger sample sizes may increase the overall accuracy. Nevertheless, we believe that the consistency in interpretation with past literature adds validity to our findings, in showing for a preliminary utility in classification of emotional content of music by listener comments. This may have potential application areas such as music therapy, where 'nostalgic' songs can potentially be categorised efficiently and used in music-based dementia interventions (Tang et al., 2018). Our research also focused on Japanese comments for Japanese songs, but future research can extend this to different cultures and languages.

References

Marios Biskas, Wing-Yee Cheung, Jacob Juhl, Constantine Sedikides, Constantine Sedikides, Tim Wildschut, and Erica Hepper. 2019. A prologue to nostalgia: savouring creates nostalgic memories that foster optimism. *Cognition and Emotion*, 33(3):417—427.

David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3(null):993–1022.

Sergio Davalos, Altaf Merchant, Gregory Rose, Brent Lessley, and Ankur Teredesia. 2015. 'the good old days': An examination of nostalgia in facebook posts. *International Journal of Human-Computer Studies*, 83:83–93.

Jerome H. Friedman. 2001. Greedy function approximation: A gradient boosting machine. *Ann. Statist.*, 29(5):1189–1232.

Brandon Greenwell, Bradley Boehmke, Jay Cunningham, and GBM Developers. 2019. *gbm: Generalized Boosted Regression Models*. R package version 2.1.5.

Brandon M. Greenwell. 2017. pdp: An r package for constructing partial dependence plots. *The R Journal*, 9(1):421–436.

Thomas Griffiths and Mark Steyvers. 2004. Finding scientific topics. *Proceedings of the National Academy of Sciences of the United States of America*, 101 Suppl 1:5228–35.

Bettina Grun and Kurt Hornik. 2011. topicmodels: An R package for fitting topic models. *Journal of Statistical Software*, 40(13):1–30.

Motohiro Ishida. 2018. *RMeCab: interface to MeCab*. R package version 1.00.

Carol Lynne Krumhansl and Justin Adam Zupnick. 2013. Cascading reminiscence bumps in popular music. *Psychological Science*, 24(10):2057–2068. PMID: 24006129.

Taku Kudo. 2005. Mecab : Yet another part-of-speech and morphological analyzer.

Max Kuhn. 2019. *caret: Classification and Regression Training*. R package version 6.0-82.

Christoph Molnar, Bernd Bischl, and Giuseppe Casalicchio. 2018. iml: An r package for interpretable machine learning. *Journal of Open Source Software*, 3(26):786.

Robert Plutchik. 1980. A general psychoevolutionary theory of emotion. In Robert Plutchik and Henry Kellerman, editors, *Theories of Emotion*, pages 3 – 33. Academic Press.

Lorenzo Porcaro and Horacio Saggion. 2019. Recognizing musical entities in user-generated content. *Computación y Sistemas*, 23.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Chelsea A. Reid, Jeffrey D. Green, Tim Wildschut, and Constantine Sedikides. 2015. Scent-evoked nostalgia. *Memory*, 23(2):157–166. PMID: 24456210.

Constantine Sedikides and Tim Wildschut. 2019. The sociality of personal and collective nostalgia. *European Review of Social Psychology*, 30(1):123–173.

Daisaku Shibata, Shoko Wakamiya, Kaoru Ito, and Eiji Aramaki. 2017. Jiwc: *Kuraudosooshinngu ni yoru nihongo kanjouhyougen jisho no kousaku* [jiwc: Construction of a japanese emotion-term dictionary through crowdsourcing]. In *23rd Annual Meeting of the Association for Natural Language Processing*, Tsukuba, Japan.

Gaurav Sood. 2019. *tuber: Access YouTube from R*. R package version 0.9.8.

Qiubi Tang, Ying Zhou, Shuixian Yang, Wong Kwok Shing Thomas, Graeme D. Smith, Zhi Yang, Lexin Yuan, and Joanne Wai yee Chung. 2018. Effect of music intervention on apathy in nursing home residents with dementia. *Geriatric Nursing*, 39(4):471 – 476.

Liila Taruffi and Stefan Koelsch. 2014. The paradox of music-evoked sadness: An online survey. *PLOS ONE*, 9(10):1–17.

Wijnand Tilburg, Martin Bruder, Tim Wildschut, Constantine Sedikides, and Anja Göritz. 2018. An appraisal profile of nostalgia. *Emotion*, 19.

Wijnand Tilburg, Tim Wildschut, and Constantine Sedikides. 2017. Nostalgia's place among self-relevant emotions. *Cognition and Emotion*, 32.

Joseph Timoney, Brian Davis, and Adarsh Raj. 2018. Nostalgic sentiment analysis of youtube comments for chart hits of the 20th century. In *AICS*.

Matthew Vess, Jamie Arndt, Clay Routledge, Constantine Sedikides, and Tim Wildschut. 2012. Nostalgia as a resource for the self. *Self and Identity*, 11.