

Other Topics You May Also Agree or Disagree: Modeling Inter-Topic Preferences using Tweets and Matrix Factorization

Akira Sasaki, Kazuaki Hanawa, Naoaki Okazaki, and Kentaro Inui

Graduate School of Information Sciences

Tohoku University

{aki-s, hanawa, okazaki, inui}@ecei.tohoku.ac.jp

Abstract

We present in this paper our approach for modeling inter-topic preferences of Twitter users: for example, *those who agree with the Trans-Pacific Partnership (TPP) also agree with free trade*. This kind of knowledge is useful not only for stance detection across multiple topics but also for various real-world applications including public opinion surveys, electoral predictions, electoral campaigns, and online debates. In order to extract users' preferences on Twitter, we design linguistic patterns in which people agree and disagree about specific topics (e.g., "A is completely wrong"). By applying these linguistic patterns to a collection of tweets, we extract statements agreeing and disagreeing with various topics. Inspired by previous work on item recommendation, we formalize the task of modeling inter-topic preferences as matrix factorization: representing users' preferences as a user-topic matrix and mapping both users and topics onto a latent feature space that abstracts the preferences. Our experimental results demonstrate both that our proposed approach is useful in predicting missing preferences of users and that the latent vector representations of topics successfully encode inter-topic preferences.

1 Introduction

Social media have changed the way people shape public opinion. The latest survey by the Pew Research Center reported that a majority of US adults (62%) obtain news via social media, and of those, 18% do so often (Gottfried and Shearer, 2016). Given that news and opinions are shared

and amplified by friend networks of individuals (Jamieson and Cappella, 2008), individuals are thereby isolated from information that does not fit well with their opinions (Pariser, 2011). Ironically, cutting-edge social media technologies promote ideological groups even with its potential to deliver diverse information.

A large number of studies already analyzed discussions, interactions, influences, and communities on social media along the political spectrum from liberal to conservative (Adamic and Glance, 2005; Zhou et al., 2011; Cohen and Ruths, 2013; Bakshy et al., 2015; Wong et al., 2016). Even though these studies provide intuitive visualizations and interpretations along the liberal-conservative axis, political analysts argue that the axis is flawed and insufficient for representing public opinion and ideologies (Kerlinger, 1984; Maddox and Lilie, 1984).

A potential solution for analyzing multiple axes of the political spectrum on social media is stance detection (Thomas et al., 2006; Somasundaran and Wiebe, 2009; Murakami and Raymond, 2010; Anand et al., 2011; Walker et al., 2012; Mohammad et al., 2016; Johnson and Goldwasser, 2016), whose task is to determine whether the author of a text is for, neutral, or against a topic (e.g., *free trade, immigration, abortion*). However, stance detection across different topics is extremely difficult. Anand et al. (2011) reported that a sophisticated method with topic-dependent features substantially improved the performance of stance detection within a topic, but such an approach could not outperform a baseline method with simple n -gram features when evaluated across topics. More recently, all participants of SemEval 2016 Task 6A (with five topics) could not outperform the baseline supervised method using n -gram features (Mohammad et al., 2016).

In addition, stance detection encounters dif-

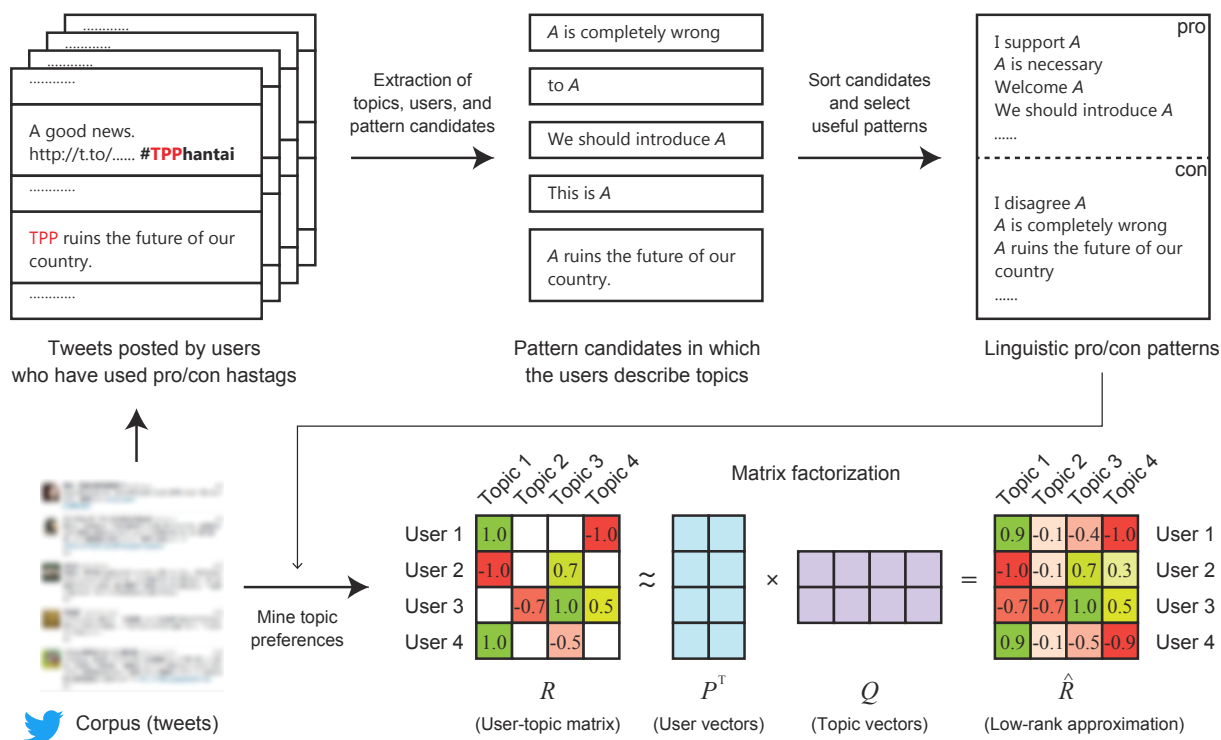


Figure 1: An overview of this study.

faculties with different user types. Cohen and Ruths (2013) observed that existing methods on stance detection fail on “ordinary” users because such methods primarily obtain training and test data from politically vocal users (e.g., politicians); for example, they found that a stance detector trained on a dataset with politicians achieved 91% accuracy on other politicians but only achieved 54% accuracy on “ordinary” users. Establishing a bridge across different topics and users remains a major challenge not only in stance detection, but also in social media analytics.

An important component in establishing this bridge is commonsense knowledge about topics. For example, consider a topic *a revision of Article 96 of the Japanese Constitution*. We infer that the statement “we should maintain armed forces” tends to favor this topic even without any lexical overlap between the topic and the statement. This inference is reasonable because: the writer of the statement favors *armed forces*; those who favor *armed forces* also favor *a revision of Article 9*¹; and those who favor *a revision of Article 9* also favor *a revision of Article 96*². In general, this kind of commonsense knowledge can be expressed in

¹Article 9 prohibits armed forces in Japan.

²Article 96 specifies high requirements for making amendments to Constitution of Japan (including Article 9).

the format: *those who agree/disagree with topic A also agree/disagree with topic B*. We call this kind of knowledge *inter-topic preference* throughout this paper.

We conjecture that previous work on stance detection indirectly learns inter-topic preferences within the same target through the use of *n*-gram features on a supervision data. In contrast, in the present paper, we directly acquire inter-topic preferences from an unlabeled corpus of tweets. This acquired knowledge regarding inter-topic preferences is useful not only for stance detection, but also for various real-world applications including public opinion survey, electoral campaigns, electoral predictions, and online debates.

Figure 1 provides an overview of this work. In our system, we extract linguistic patterns in which people agree and disagree about specific topics (e.g., “A is completely wrong”); to accomplish this, as described in Section 2.1, we make use of hashtags within a large collection of tweets. The patterns are then used to extract instances of users’ preferences regarding various topics, as detailed in Section 2.2. Inspired by previous work on item recommendation, in Section 3, we formalize the task of modeling inter-topic preferences as a matrix factorization: representing a sparse user-topic matrix (i.e., the extracted instances) with the prod-

uct of low-rank user and topic matrices. These low-rank matrices provide *latent vector representations* of both users and topics. This approach is also useful for completing preferences of “ordinary” (i.e., less vocal) users, which fills the gap between different types of users.

The contributions of this paper are threefold.

1. To the best of our knowledge, this is the first study that models inter-topic preferences for unlimited targets on real-world data.
2. Our experimental results show that this approach can accurately predict missing topic preferences of users accurately (80–94%).
3. Our experimental results also demonstrate that the latent vector representations of topics successfully encode inter-topic preferences, e.g., *those who agree with nuclear power plants also agree with nuclear fuel cycles*.

This study uses a Japanese Twitter corpus because of its availability from the authors, but the core idea is applicable to any language.

2 Mining Topic Preferences of Users

In this section, we describe how we collect statements in which users agree or disagree with various topics on Twitter, which then serves as source data for modeling inter-topic preferences. More formally, we are interested in acquiring a collection of tuples (u, t, v) , where: $u \in U$ is a user; U is the set of all users on Twitter; $t \in T$ is a topic; T is the set of all topics; and $v \in \{+1, -1\}$ is $+1$ when the user u agrees with the topic t and -1 otherwise (i.e., disagreement).

Throughout this work, we use a corpus consisting of 35,328,745,115 Japanese tweets (7,340,730 users) crawled from February 6, 2013 to September 30, 2016. We removed retweets from the corpus.

2.1 Mining Linguistic Patterns of Agreement and Disagreement

We use linguistic patterns to extract tuples (u, t, v) from the aforementioned corpus. More specifically, when a tweet message matches to one of linguistic patterns of agreement (e.g., “ t is necessary”), we regard that the author u of the tweet agrees with topic t . Conversely, a statement of disagreement is identified by linguistic patterns for disagreement (e.g., “ t is unacceptable”).

In order to design linguistic patterns, we focus on hashtags appearing in the corpus that have been popular clues for locating subjective statements such as sentiments (Davidov et al., 2010), emotions (Qadir and Riloff, 2014), and ironies (Van Hee et al., 2016). Hashtags are also useful for finding strong supporters and critics, as well as their target topics; for example, #immigrantsWelcome indicates that the author favors *immigrants*; and #StopAbortion is against *abortion*.

Based on this intuition, we design regular expressions for both *pro hashtags* “#(.+)sansei”³ and *con hashtags* “#(.+)hantai”⁴, where $(.+)$ matches a target topic. These regular expressions can find users who have strong preferences to topics. Using this approach, we extracted 31,068 occurrences of pro/con hashtags used by 18,582 users for 4,899 topics. We regard the set of topics found using this procedure as set of target topics T in this study.

Each time we encounter a tweet containing a pro/con hashtag, we searched for corresponding textual statements as follows. Suppose that a tweet includes a hashtag (e.g., #TPPsansei) for a topic t (e.g., *TPP*). Assuming that the author of the given tweet does not change their attitude toward a topic over time, we search for other tweets posted by the same author that also have the topic keyword t . This process retrieves tweets like “I support TPP.” Then, we replace the topic keyword into a variable A to extract patterns, e.g., “I support A .” Here, the definition of the pattern unit is language specific. For Japanese tweets, we simply recognize a pattern that starts with a variable (i.e., topic) and ends at the end of the sentence⁵.

Because this procedure also extracts useless patterns such as “to A ” and “this is A ”, we manually choose useful patterns in a systematic way: sort patterns in descending order of the number of users who use the pattern; and check the sorted list of patterns manually; and remove useless patterns.

³Unlike English hashtags, we systematically attach a noun *sansei*, which stands for *pro* (agreement) in Japanese, to a topic, for example, #TPPsansei. This paper uses the alphabetical expression *sansei* only for explanation; the actual pattern uses Chinese characters corresponding to *sansei*.

⁴A Japanese noun *hantai* stands for *con* (disagreement), for example, #TPPhantai. This paper uses the alphabetical expression *hantai* only for explanation; the actual pattern uses Chinese characters corresponding to *hantai*.

⁵In English, this treatment roughly corresponds to extracting a verb phrase with the variable A .

Using this approach, we obtained 100 pro patterns (e.g., “welcome A ” and “ A is necessary”) and 100 con patterns (“do not let A ” and “I don’t want A ”).

2.2 Extracting Instances of Topic Preferences

By using the pro and con patterns acquired using the approach described in Section 2.1, we extract instances of (u, t, v) as follows. When a sentence in a tweet whose author is user u matches one of the pro patterns (e.g., “ t is necessary”) and the topic t is included in the set of target topics T , we recognize this as an instance of $(u, t, +1)$. Similarly, when a sentence matches one of the con patterns (e.g., “I don’t want t ”) and the topic t is included in the set of target topics T , we recognize this as an instance of $(u, t, -1)$. Using this approach, we collected 25,805,909 tuples corresponding to 3,302,613 users and 4,899 topics. Because these collected tuples included comparatively infrequent users and topics, we removed users and topics that appeared less than five times. In addition, there were also meaningless frequent topics such as “of” and “it”. Therefore, we sorted topics in descending order of their co-occurrence frequencies with each of the pro patterns and con patterns, and then removed meaningless topics in the top 100 topics. This resulted in 9,961,509 tuples regarding 273,417 users and 2,323 topics.

3 Matrix Factorization

Using the methods described in Section 2, from the corpus, we collected a number of instances of users’ preferences regarding various topics. However, Twitter users do not necessarily express preferences for all topics. In addition, it is by nature impossible to predict whether a new (i.e., nonexistent in the data) user agrees or disagrees with given topics. Therefore, in this section, we apply matrix factorization (Koren et al., 2009) in order to predict missing values, inspired by research regarding item recommendation (Bell and Koren, 2007; Dror et al., 2011). In essence, matrix factorization maps both users and topics onto a latent feature space that abstracts topic preferences of users.

Here, let R be a sparse matrix of $|U| \times |T|$. Only when a user u expresses a preference for topic t do we compute an element of the sparse matrix $r_{u,t}$,

$$r_{u,t} = \frac{\#(u, t, +1) - \#(u, t, -1)}{\#(u, t, +1) + \#(u, t, -1)} \quad (1)$$

Here, $\#(u, t, +1)$ and $\#(u, t, -1)$ represent the numbers of occurrences of instances $(u, t, +1)$

and $(u, t, -1)$, respectively. Thus, an element $r_{u,t}$ approaches $+1$ as the user u favors the topic t , and -1 otherwise. If the user u does not make any statement regarding the topic t (i.e., neither $(u, t, +1)$ nor $(u, t, -1)$ exists in the data), we do not fill the corresponding element, leaving it as a missing value.

Matrix factorization decomposes the sparse matrix R into low-dimensional matrices $P \in \mathbb{R}^{k \times |U|}$ and $Q \in \mathbb{R}^{k \times |T|}$, where k is a parameter that specifies the number of dimensions of the latent space. We minimize the following objective function to find the matrices P and Q ,

$$\min_{P, Q} \sum_{(u,t) \in R} \left((r_{u,t} - \mathbf{p}_u^\top \mathbf{q}_t)^2 + \lambda_P \|\mathbf{p}_u\|^2 + \lambda_Q \|\mathbf{q}_t\|^2 \right). \quad (2)$$

Here, $(u, t) \in R$ is repeated for elements filled in the sparse matrix R , $\mathbf{p}_u \in \mathbb{R}^k$ and $\mathbf{q}_t \in \mathbb{R}^k$ are u column vectors of P and v column vectors of Q , respectively, and $\lambda_P \geq 0$ and $\lambda_Q \geq 0$ represent coefficients of regularization terms. We call \mathbf{p}_u and \mathbf{q}_t the *user vector* and *topic vector*, respectively.

Using these user and topic vectors, we can predict an element $\hat{r}_{u,t}$ that may be missing in the original matrix R ,

$$\hat{r}_{u,t} \simeq \mathbf{p}_u^\top \mathbf{q}_t. \quad (3)$$

We use `libmf`⁶ (Chin et al., 2015) to solve the optimization problem in Equation 2. We set regularization coefficients $\lambda_P = 0.1$ and $\lambda_Q = 0.1$ and use default values for the other parameters of `libmf`.

4 Evaluation

4.1 Determining the Dimension Parameter k

How good is the low-rank approximation found by matrix factorization? And can we find the “sweet spot” for the number of dimensions k of the latent space? We investigate the reconstruction error of matrix factorization using different values of k to answer these questions. We use Root Mean Squared Error (RMSE) to measure error,

$$RMSE = \sqrt{\frac{\sum_{(u,t) \in R} (\mathbf{p}_u^\top \mathbf{q}_t - r_{u,t})^2}{N}}. \quad (4)$$

⁶<https://github.com/cjlin1/libmf>

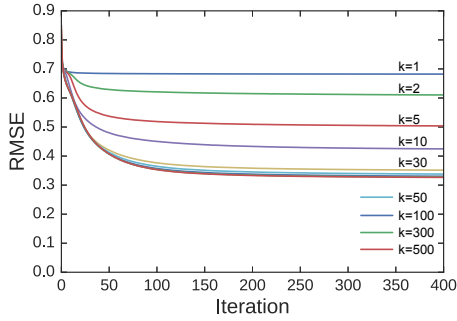


Figure 2: Reconstruction error (RMSE) of matrix factorization with different k .

Here, N is the number of elements in the sparse matrix R (i.e., the number of known values).

Figure 2 shows RMSE values over iterations of `libmf` with the dimension parameter $k \in \{1, 2, 5, 10, 30, 50, 100, 300, 500\}$. We observed that the reconstruction error decreased as the iterative method of `libmf` progressed. The larger the number of dimensions k was, the smaller the reconstruction error became; the lowest reconstruction error was 0.3256 with $k = 500$. We also observed the error with $k = 1$, which corresponds to mapping users and topics onto one dimension similarly to the political spectrum of liberal and conservative. Judging from the relatively high RMSE values with $k = 1$, we conclude that it may be difficult to represent everything in the data using a one-dimensional axis. Based on this result, we concluded that matrix factorization with $k = 100$ is sufficient for reconstructing the original matrix R and therefore used this parameter value for the rest of our experiments.

4.2 Predicting Missing Topic Preferences

How accurately can the user and topic vectors predict missing topic preferences? To answer this question, we evaluate the accuracy in predicting hidden preferences in the matrix R as follows. First, we randomly selected 5% of existing elements in R and let Y represent the collection of the selected elements (test set). We then perform matrix factorization on the sparse matrix without the selected elements of Y , that is, only with the remaining 95% elements of R (training set). We define the accuracy of the prediction as

$$\frac{1}{|Y|} \sum_{u,t \in Y} \mathbb{1}(\text{sign}(\hat{r}_{u,t}) = \text{sign}(r_{u,t})) \quad (5)$$

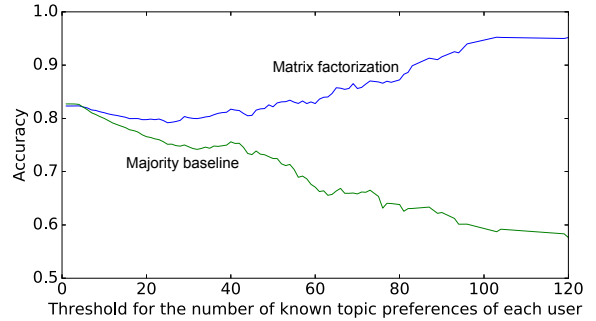


Figure 3: Prediction accuracy when changing the threshold for the number of known topic preferences of each user.

Here, $r_{u,t}$ denotes the actual (i.e., self-declared) preference values, $\hat{r}_{u,t}$ represents the preference value predicted by Equation 3, $\text{sign}(\cdot)$ represents the sign of the argument, and $\mathbb{1}(\cdot)$ yields 1 only when the condition described in the argument holds and 0 otherwise. In other words, Equation 5 computes the proportion of correct predictions to all predictions, assuming zero to be the decision boundary between pro and con.

Figure 3 plots prediction accuracy values calculated from different sets of users. Here the x-axis represents a threshold θ , which filters out users whose declarations of topic preferences are no greater than θ topics. In other words, Figure 3 shows prediction accuracy when we know user preferences for at least θ topics. For comparison, we also include the majority baseline that predicts pro and con based on the majority of preferences regarding each topic in the training set.

Our proposed method was able to predict missing preferences with an 82.1% accuracy for users stating preferences for at least five topics. This accuracy increased as our method received more information regarding the users, reaching a 94.0% accuracy when $\theta = 100$. This result again indicates that our proposed method reasonably utilizes known preferences to complete missing preferences.

In contrast, the performance of the majority baseline decreased as it received more information regarding the users. Because this result was rather counter-intuitive, we examined the cause of this phenomenon. Consequently, this result turned out to be reasonable because preferences of vocal users deviated from those of the average users. Figure 4 illustrates this finding, showing the mean of variances of preference values $r_{u,t}$ across self-

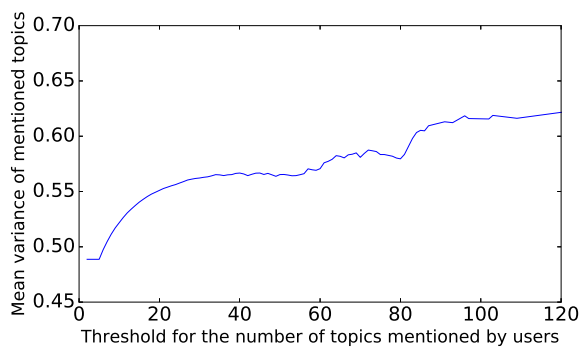


Figure 4: Mean variance of preference values of self-declared topics when changing the threshold for the number of self-declared topics.

declared topics. In the figure, the x-axis represents a threshold θ , which filters out users whose statements of topic preferences are no greater than θ topics. We observe that the mean variance increased as we focused on vocal users. Overall, these results demonstrate the usefulness of user and topic vectors in predicting missing preferences.

Table 1 shows examples in which missing preferences of two users were predicted from known statements of agreements and disagreements⁷. In the table, predicted topics are accompanied by the corresponding $\hat{r}_{u,t}$ value in parentheses. As an example, our proposed method predicted that the user A, who is positive toward *regime change* but negative toward *Okinawa US military base*, may also be positive toward *vote of non-confidence to Cabinet* but negative toward *construction of a new base*.

4.3 Inter-topic Preferences

Do the topic vectors obtained by matrix factorization capture inter-topic preferences, such as “People who agree with A also agree with B”?

Because no dataset exists for this evaluation, we created a dataset of pairwise inter-topic preferences by using a crowdsourcing service⁸. Sampling topic pairs randomly, we collected 150 topic pairs whose cosine similarities of topic vectors

⁷We anonymized user names in these examples. In addition, we removed topics that are too discriminatory or aggressive to other countries and races. Even though the experimental results of this paper do not necessarily reflect our idea, we do not think it is a good idea to distribute politically incorrect ideas through this paper.

⁸We used Yahoo! Crowdsourcing, a Japanese online service for crowdsourcing.
<http://crowdsourcing.yahoo.co.jp/>

were below -0.6 , 150 pairs whose cosine similarities were between -0.6 and 0.6 , and 150 pairs whose cosine similarities were above 0.6 . In this way, we obtained 450 topic pairs for evaluation.

Given a pair of topics A and B , a crowd worker was asked to choose a label from the following three options: (a) *those who agree/disagree with topic A may also agree/disagree with topic B*; (b) *those who agree/disagree with topic A may conversely disagree/agree with topic B*; (c) *otherwise (no association between A and B)*. Creating twenty pairs of topics as gold data, we removed labeling results from workers whose accuracy is less than 90%.

Consequently, we obtained 6–10 human judgements for every topic pair. Regarding (a) as +1 point, (b) as -1 point, and (c) as 0 point, we computed the mean of the points (i.e., average human judgements) for each topic pair. Spearman’s rank correlation coefficient (ρ) between cosine similarity values of topic vectors and human judgements was 0.2210. We could observe a moderate correlation even though inter-topic preferences collected in this manner were highly subjective.

In addition to the quantitative evaluation, as summarized in Table 2, we also checked similar topics for three controversial topics, *Liberal Democratic Party (LDP)*, *constitutional amendment* and *right of foreigners to vote* (Table 2). Topics similar to LDP included synonymous ones (e.g., *Abe’s LDP* and *Abe administration*) and other topics promoted by the LDP (e.g., *resuming nuclear power plant operations*, *bus rapid transit (BRT)* and *hate speech countermeasure law*). Considering that people who support the LDP may also tend to favor its policies, we found these results reasonable. As for the other example, *constitutional amendment* had a feature vector that was similar to that of *amendment of Article 9*, *enforcement of specific secret protection law* and *security related law*. From these results, we concluded that topic vectors were able to capture inter-topic preferences.

5 Related Work

In this section, we summarize the related work that spreads across various research fields.

Social Science and Political Science A number of studies analyze social phenomena regarding political activities, political thoughts, and public opinions on social media. These studies

| User | Type | Topic |
|------|--------------------------|---|
| A | Agreement (declared) | regime change, capital relocation |
| | Disagreement (declared) | Okinawa US military base, nuclear weapons, TPP, Abe Cabinet, Abe government, nuclear cycle, right to collective defense, nuclear power plant, Abenomics |
| | Agreement (predicted) | same-sex partnership ordinance (0.9697), vote of non-confidence to Cabinet (0.9248), national people’s government (0.9157), abolition of tax (0.8978) |
| | Disagreement (predicted) | steamrolling war bill (-1.0522), worsening dispatch law (-1.0301), Sendai nuclear power plant (-1.0269), war bill (-1.0190), construction of a new base (-1.0186), Abe administration (-1.0173), landfill Henoko (-1.0158), unreasonable arrest (-1.0113) |
| B | Agreement (declared) | visit shrine, marriage |
| | Disagreement(declared) | tax increase, conscription, amend Article 9 |
| | Agreement (predicted) | national people’s government (0.8467), abolition of tax (0.8300), same-sex partnership ordinance (0.7700), security bills (0.6736) |
| | Disagreement (predicted) | corporate tax cuts (-1.0439), Liberal Democratic Party’s draft constitution (-1.0396), radioactivity (-1.0276), rubble (-1.0159), nuclear cycle (-1.0143) |

Table 1: Examples of agreement/disagreement topics predicted for two sample users A and B, with predicted score $\hat{r}_{u,v}$ shown in parenthesis.

| Topic | Topics with a high degree of cosine similarity |
|--------------------------------|---|
| Liberal Democratic Party (LDP) | Abe’s LDP (0.3937), resuming nuclear power plant operations (0.3765), bus rapid transit (BRT) (0.3410), hate speech countermeasure law (0.3373), Henoko relocation (0.3353), C-130 (0.3338), Abe administration (0.3248), LDP & Komeito (0.2898), Prime Minister Abe (0.2835) |
| constitutional amendment | amendment of Article 9 (0.4520), enforcement of specific secret protection law (0.4399), security related law (0.4242), specific confidentiality protection law (0.4022), security bill amendment (0.3977), defense forces (0.3962), my number law (0.3874), collective self-defense rights (0.3687), militarist revival (0.3567) |
| right of foreigners to vote | human rights law (0.5405), anti-discrimination law (0.5376), hate speech countermeasure law (0.5080), foreigner’s life protection (0.4553), immigration refugee (0.4520), co-organized Olympics (0.4379) |

Table 2: Topics identified as being similar to the three controversial topics shown in the left column.

model the political spectrum from liberal to conservative (Adamic and Glance, 2005; Zhou et al., 2011; Cohen and Ruths, 2013; Bakshy et al., 2015; Wong et al., 2016), political parties (Tumasjan et al., 2010; Boutet et al., 2013; Makazhanov and Rafiei, 2013), and elections (O’Connor et al., 2010; Conover et al., 2011).

Employing a single axis (e.g., liberal to conservative) or a few axes (e.g., political parties and candidates of elections), these studies provide intuitive visualizations and interpretations along the respective axes. In contrast, this study is the first attempt to recognize and organize various axes of topics on social media with no prior assumptions regarding the axes. Therefore, we think our study provides a new tool for computational social science and political science that enables researchers to analyze and interpret phenomena on social media.

Next, we describe previous research focused on acquiring lexical knowledge of politics. Sim et al. (2013) measured ideological positions of candidates in US presidential elections from their

speeches. The study first constructs “cue lexicons” from political writings labeled with ideologies by domain experts, using sparse additive generative models (Eisenstein et al., 2011). These constructed cue lexicons were associated with such ideologies as *left*, *center*, and *right*. Representing each speech of a candidate with cue lexicons, they inferred the proportions of ideologies of the candidate. The study requires a predefined set of labels and text data associated with the labels.

Bamman and Smith (2015) presented an unsupervised method for assessing the political stance of a proposition, such as “global warming is a hoax,” along the political spectrum of liberal to conservative. In their work, a proposition was represented by a tuple in the form ⟨subject, predicate⟩, for example, ⟨*global warming*, *hoax*⟩. They presented a generative model for users, subjects, and predicates to find a one-dimensional latent space that corresponded to the political spectrum.

Similar to our present work, their work (Bamman and Smith, 2015) did not require labeled data

to map users and topics (i.e., subjects) onto a latent feature space. In their paper, they reported that the generative model outperformed Principal Component Analysis (PCA), which is a method for matrix factorization. Empirical results here probably reflected the underlying assumptions that PCA treats missing elements as zero and not as missing data. In contrast, in the present work, we properly distinguish missing values from zero, excluding missing elements of the original matrix from the objective function of Equation 2. Further, this work demonstrated the usefulness of the latent space, that is, topic and user vectors, in predicting missing topic preferences of users and inter-topic preferences.

Fine-grained Opinion Analysis The method presented in Section 2 is an instance of fine-grained opinion analysis (Wiebe et al., 2005; Choi et al., 2006; Johansson and Moschitti, 2010; Yang and Cardie, 2013; Deng and Wiebe, 2015), which extracts a tuple of a subjective opinion, a holder of the opinion, and a target of the opinion from text. Although these previous studies have the potential to improve the quality of the user-topic matrix R , unfortunately, no corpus or resource is available for the Japanese language. We do not currently have a large collection of English tweets, but combining fine-grained opinion analysis with matrix factorization is an immediate future work.

Causality Relation Some of inter-topic preferences in this work can be explained by causality relation, for example, “TPP promotes free trade.” A number of previous studies acquire instances of causal relation (Girju, 2003; Do et al., 2011) and promote/suppress relation (Hashimoto et al., 2012; Fluck et al., 2015) from text. The causality knowledge is useful for predicting (hypotheses of) future events (Radinsky et al., 2012; Radinsky and Davidovich, 2012; Hashimoto et al., 2015).

Inter-topic preferences, however, also include pairs of topics in which causality relation hardly holds. As an example, it is unreasonable to infer that *nuclear plant* and *railroading of bills* have a causal relation, but those who dislike *nuclear plant* also oppose *railroading of bills* because presumably they think the governing political parties rush the bill for resuming a nuclear plant. In this study, we model these inter-topic preferences based on preferences of the public. That said, we have as a promising future direction of our work plans to in-

corporate approaches to acquire causality knowledge.

6 Conclusion

In this paper, we presented a novel approach for modeling inter-topic preferences of users on Twitter. Designing linguistic patterns for identifying support and opposition statements, we extracted users’ preferences regarding various topics from a large collection of tweets. We formalized the task of modeling inter-topic preferences as a matrix factorization that maps both users and topics onto a latent feature space that abstracts users’ preferences. Through our experimental results, we demonstrated that our approach was able to accurately predict missing topic preferences of users (80–94%) and that our latent vector representations of topics properly encoded inter-topic preferences.

For our immediate future work, we plan to embed the topic and user vectors to create a cross-topic stance detector. It is possible to generalize our work to model heterogeneous signals, such as interests and behaviors of people, for example, “those who are interested in A also support B ,” and “those who favor A also vote for B ”. Therefore, we believe that our work will bring about new applications in the field of NLP and other disciplines.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 15H05318 and JST CREST Grant Number J130002054, Japan.

References

- Lada A. Adamic and Natalie Glance. 2005. [The political blogosphere and the 2004 U.S. election: Divided they blog.](#) In *Proceedings of the 3rd International Workshop on Link Discovery (LinkKDD 2005)*. pages 36–43. <https://doi.org/10.1145/1134271.1134277>.
- Pranav Anand, Marilyn Walker, Rob Abbott, Jean E. Fox Tree, Robeson Bowmani, and Michael Minor. 2011. [Cats rule and dogs drool!: Classifying stance in online debate.](#) In *Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (WASSA 2011)*. pages 1–9.
- Eytan Bakshy, Solomon Messing, and Lada A. Adamic. 2015. [Exposure to ideologically diverse news and opinion on face-](#)

- book. *Science* 348(6239):1130–1132. <https://doi.org/10.1126/science.aaa1160>.
- David Bamman and Noah A. Smith. 2015. Open extraction of fine-grained political statements. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP 2015)*. pages 76–85. <https://doi.org/10.18653/v1/D15-1008>.
- Robert M. Bell and Yehuda Koren. 2007. Lessons from the Netflix prize challenge. *ACM SIGKDD Explorations Newsletter* 9(2):75–79. <https://doi.org/10.1145/1345448.1345465>.
- Antoine Boutet, Hyoungshick Kim, and Eiko Yoneki. 2013. What’s in Twitter, I know what parties are popular and who you are supporting now! *Social Network Analysis and Mining (SNAM 2012)* 3(4):1379–1391. <https://doi.org/10.1109/ASONAM.2012.32>.
- Wei-Sheng Chin, Yong Zhuang, Yu-Chin Juan, and Chih-Jen Lin. 2015. A fast parallel stochastic gradient method for matrix factorization in shared memory systems. *ACM Transactions on Intelligent Systems and Technology (TIST)* 6(1):2. <https://doi.org/10.1145/2668133>.
- Yejin Choi, Eric Breck, and Claire Cardie. 2006. Joint extraction of entities and relations for opinion recognition. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP 2006)*. pages 431–439. <http://aclweb.org/anthology/W06-1651>.
- Raviv Cohen and Derek Ruths. 2013. Classifying political orientation on Twitter: It’s not easy! In *Proc. of the Seventh International AAAI Conference on Weblogs and Social Media (ICWSM 2013)*. pages 91–99.
- Michael D Conover, Bruno Gonçalves, Jacob Ratkiewicz, Alessandro Flammini, and Filippo Menczer. 2011. Predicting the political alignment of twitter users. In *Privacy, 2011 IEEE Third International Conference on Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing (PASSAT-SocialCom 2011)*. IEEE, pages 192–199.
- Dmitry Davidov, Oren Tsur, and Ari Rappoport. 2010. Enhanced sentiment learning using twitter hashtags and smileys. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*. pages 241–249. <http://aclweb.org/anthology/C10-2028>.
- Lingjia Deng and Janyce Wiebe. 2015. MPQA 3.0: An entity/event-level sentiment corpus. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2015)*. pages 1323–1328. <https://doi.org/10.3115/v1/N15-1146>.
- Quang Do, Yee Seng Chan, and Dan Roth. 2011. Minimally supervised event causality identification. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP 2011)*. pages 294–303. <http://aclweb.org/anthology/D11-1027>.
- Gideon Dror, Noam Koenigstein, Yehuda Koren, and Markus Weimer. 2011. The Yahoo! Music dataset and KDD-Cup’11. In *Proceedings of the 2011 International Conference on KDD Cup 2011 (KDDCUP 2011)*. pages 3–18.
- Jacob Eisenstein, Amr Ahmed, and Eric P Xing. 2011. Sparse additive generative models of text. In *Proceedings of the 28th International Conference on Machine Learning (ICML 2011)*.
- Juliane Fluck, Sumit Madan, Tilia Renate Ellendorff, Theo Mevissen, Simon Clematide, Adrian van der Lek, and Fabio Rinaldi. 2015. Track 4 overview: Extraction of causal network information in biological expression language (BEL). In *Proceedings of the Fifth BioCreative Challenge Evaluation Workshop*. pages 333–346.
- Roxana Girju. 2003. Automatic detection of causal relations for question answering. In *Proceedings of the ACL 2003 Workshop on Multilingual Summarization and Question Answering - Volume 12*. pages 76–83. <https://doi.org/10.3115/1119312.1119322>.
- Jeffrey Gottfried and Elisa Shearer. 2016. News use across social media platforms 2016. Technical report, Pew Research Center.
- Chikara Hashimoto, Kentaro Torisawa, Stijn De Saeger, Jong-Hoon Oh, and Jun’ichi Kazama. 2012. Excitatory or inhibitory: A new semantic orientation extracts contradiction and causality from the web. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2012)*. Association for Computational Linguistics, pages 619–630. <http://aclweb.org/anthology/D12-1057>.
- Chikara Hashimoto, Kentaro Torisawa, Julien Kloetzer, and Jong-Hoon Oh. 2015. Generating event causality hypotheses through semantic relations. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI 2015)*. pages 2396–2403.
- Kathleen Hall Jamieson and Joseph N. Cappella. 2008. *Echo Chamber: Rush Limbaugh and the Conservative Media Establishment*. Oxford University Press.
- Richard Johansson and Alessandro Moschitti. 2010. Syntactic and semantic structure for opinion expression detection. In *Proceedings of the Fourteenth Conference on Computational Natural Language Learning (CoNLL 2010)*. pages 67–76. <http://aclweb.org/anthology/W10-2910>.

- Kristen Johnson and Dan Goldwasser. 2016. “All I know about politics is what I read in Twitter”: Weakly supervised models for extracting politicians’ stances from twitter. In *Proceedings of the 26th International Conference on Computational Linguistics (COLING 2016)*. pages 2966–2977. <http://aclweb.org/anthology/C16-1279>.
- Fred N. Kerlinger. 1984. *Liberalism and Conservatism: The Nature and Structure of Social Attitudes*. Lawrence Erlbaum Associates.
- Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer* 42(8):30–37. <https://doi.org/10.1109/MC.2009.263>.
- William S. Maddox and Stuart A. Lilie. 1984. *Beyond Liberal and Conservative: Reassessing the Political Spectrum*. Cato Inst.
- Aibek Makazhanov and Davood Rafiei. 2013. Predicting political preference of twitter users. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013)*. pages 298–305. <https://doi.org/10.1145/2492517.2492527>.
- Saif Mohammad, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu, and Colin Cherry. 2016. Semeval-2016 task 6: Detecting stance in tweets. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*. pages 31–41. <https://doi.org/10.18653/v1/S16-1003>.
- Akiko Murakami and Rudy Raymond. 2010. Support or oppose?: classifying positions in online debates from reply activities and opinion expressions. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*. pages 869–875. <http://aclweb.org/anthology/C10-2100>.
- Brendan O’Connor, Ramnath Balasubramanyan, Bryan R. Routledge, and Noah A. Smith. 2010. From tweets to polls: Linking text sentiment to public opinion time series. In *Proceedings of the Fourth International AAI Conference on Weblogs and Social Media (ICWSM 2010)*. pages 122–129.
- Eli Pariser. 2011. *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*. Penguin Books.
- Ashequl Qadir and Ellen Riloff. 2014. Learning emotion indicators from tweets: Hashtags, hashtag patterns, and phrases. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP 2014)*. pages 1203–1209. <https://doi.org/10.3115/v1/D14-1127>.
- Kira Radinsky and Sagie Davidovich. 2012. Learning to predict from textual data. *Journal of Artificial Intelligence Research (JAIR)* 45(1):641–684.
- Kira Radinsky, Sagie Davidovich, and Shaul Markovitch. 2012. Learning causality for news events prediction. In *Proceedings of the 21st International Conference on World Wide Web (WWW 2012)*. pages 909–918. <https://doi.org/10.1145/2187836.2187958>.
- Yanchuan Sim, Brice D. L. Acree, Justin H. Gross, and Noah A. Smith. 2013. Measuring ideological proportions in political speeches. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP 2013)*. pages 91–101. <http://aclweb.org/anthology/D13-1010>.
- Swapna Somasundaran and Janyce Wiebe. 2009. Recognizing stances in online debates. In *Joint conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing (ACL-IJCNLP 2009)*. pages 226–234. <http://aclweb.org/anthology/P09-1026>.
- Matt Thomas, Bo Pang, and Lillian Lee. 2006. Get out the vote: Determining support or opposition from congressional floor-debate transcripts. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP 2006)*. pages 327–335. <http://aclweb.org/anthology/W06-1639>.
- Andranik Tumasjan, Timm Oliver Sprenger, Philipp G Sandner, and Isabell M Welp. 2010. Predicting elections with twitter: What 140 characters reveal about political sentiment. In *Fourth International AAI Conference on Weblogs and Social Media (ICWSM 2010)*. pages 178–185.
- Cynthia Van Hee, Els Lefever, and Veronique Hoste. 2016. Monday mornings are my fave :) #not exploring the automatic recognition of irony in english tweets. In *Proceedings of the 26th International Conference on Computational Linguistics (COLING 2016)*. pages 2730–2739. <http://aclweb.org/anthology/C16-1257>.
- Marilyn A. Walker, Pranav Anand, Rob Abbott, Jean E. Fox Tree, Craig Martell, and Joseph King. 2012. That is your evidence?: Classifying stance in online political debate. *Decision Support Systems* 53(4):719–729. <https://doi.org/10.1016/j.dss.2012.05.032>.
- Janyce Wiebe, Theresa Wilson, and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation* 39(2):165–210. <https://doi.org/10.1007/s10579-005-7880-9>.
- Felix Ming Fai Wong, Chee Wei Tan, Soumya Sen, and Mung Chiang. 2016. Quantifying political leaning from tweets, retweets, and retweeters. *IEEE Transactions on Knowledge and Data Engineering* 28(8):2158–2172. <https://doi.org/10.1109/TKDE.2016.2553667>.

Bishan Yang and Claire Cardie. 2013. [Joint inference for fine-grained opinion extraction](#). In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL 2013)*, pages 1640–1649. <http://aclweb.org/anthology/P13-1161>.

Daniel Xiaodan Zhou, Paul Resnick, and Qiaozhu Mei. 2011. Classifying the political leaning of news articles and users from user votes. In *Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011)*, pages 417–424.