Abstract Venue Concept Detection from Location-Based Social Networks

Yi Liao[1(✉)], Shoaib Jameel[2], Wai Lam[1], and Xing Xie[3]

[1] Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Hong Kong, China
{yliao,wlam}@se.cuhk.edu.hk
[2] School of Computer Science and Informatics, Cardiff University, Cardiff, Wales
jameels1@cardiff.ac.uk
[3] Microsoft Research, Beijing, China
xingx@microsoft.com

Abstract. We investigate a new graphical model that can generate latent abstract concepts of venues, or Point of Interest (POI) by exploiting text data in venue profiles obtained from location-based social networks (LBSNs). Our model offers tailor-made modeling for two different types of text data that commonly appears in venue profiles, namely, tags and comments. Such modeling can effectively exploit their different characteristics. Meanwhile, the modeling of these two parts are tied with each other in a coordinated manner. Experimental results show that our model can generate better abstract venue concepts than comparative models.

Keywords: Location-based social networks · Abstract venue concept · Graphical model

1 Introduction

With the advent of online social networks such as Facebook, Foursquare, etc., geo-tagging has become a popular activity online where people broadcast their location [1,2]. Such geo-tagged information can be useful to advertisers who want to recommend a venue or a product based on the user's past movement patterns or construct more interesting lifestyle patterns as proposed in [3], where rather than estimating lifestyles from the check-in data directly, we can first convert each check-in to an abstract data, in which the specific name of the venue can be replaced by an abstract name. This will help mitigate sparsity problem to a large

The work described in this paper is substantially supported by grants from the Research Grant Council of the Hong Kong Special Administrative Region, China (Project Codes: 413510 and 14203414) and the Microsoft Research Asia Urban Informatics Grant FY14-RES-Sponsor-057. This work is also affiliated with the CUHK MoE-Microsoft Key Laboratory of Human-centric Computing and Interface Technologies.

G. Zuccon et al. (Eds.): AIRS 2015, LNCS 9460, pp. 147–157, 2015.
DOI: 10.1007/978-3-319-28940-3_12

extent. We present a novel generative probabilistic model that exploits textual information obtained from location-based social networks (LBSNs) to uncover meaningful latent concepts related to venues, or Point of Interest (POI). We call such automatically discovered concepts "abstract venue concepts". For example, a concept representing "upscale hotels" may be discovered and it may contain representative terms such as "five-star", "luxury", "expensive", etc. Abstract venue concepts enable semantic characterization of venues, facilitating a better understanding of venues for both users and service providers, which could potentially benefit services such as venue recommendation [4]. While we could use the categories provided by Foursquare and similar services, the taxonomies which are used by these LBSNs are not always sufficiently fine-grained. For example, by investigating the category tree[1] of Foursquare, we can easily observe that the LBSN assigns all types of hotels the to same category *hotel*, rather than distinguishing finer properties of the hotel, such as "upscale hotel". Moreover, these taxonomies are LBSN-specific, which causes problems when we want to integrate check-in data from different LBSNs.

Text data obtained from LBSNs has also been used in geographical discovery such as [5–7]. Kim et al. [8] recently applied Latent Dirichlet Allocation (`LDA`) [9] to elicit the semantic concepts of venues with aggregated text data from venue profiles. However, we observe that the text data in venue profiles originates from two different sources. The first source is *tags*, which is a set of discrete terms describing the intrinsic properties of the venue, e.g. "hotels", "shopping mall", etc. Tags are usually drawn from a relatively fixed word lexicon. The second source is *comments*, which are sentences written by users expressing their opinions about the venue. The linguistic property of comments is rather different from tags in that it consists of natural expressions which can be grammatical or ungrammatical, written by any users. Table 1 shows the venue document of the Ritz-Carlton, an upscale hotel located in Hong Kong. It contains the whole tag set and an example of comments.

One novelty of our model is that we consider tags and comments separately, and our proposed model offers tailor-made modeling for these two kinds of text data, exploiting their different characteristics. Meanwhile, the modeling of these two parts are tied to each other in a coordinated manner which makes our approach considerably different from existing approaches. Experimental results obtained by our model are more superior than other comparative models.

Table 1. Venue document of the Ritz-Carlton, Hong Kong

Types	Value
tags	hotel, five-star, icc, international commerce centre, luxury
comments	"Amazing stay. Gorgeous design in every detail. Guests enjoy panoramic view of Hong Kong from all corners. Stunning views + over the top design. Great staffs!"

[1] https://developer.foursquare.com/categorytree.

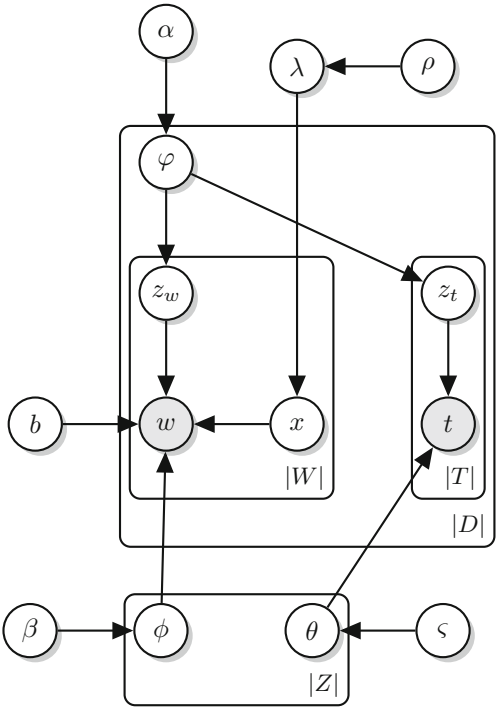

Fig. 1. The graphical model for abstract venue detection.

2 Abstract Venue Concept Detection

For each particular venue on a LBSN such as Foursquare, textual information, namely tags and user comments, are aggregated into a single document, called venue profile document. The category names of the given venue are also included as tags in our model. As a result, a venue profile document is composed of a bag of *tags* and a bag of *words* extracted from user comments.

The graphical model for our proposed abstract venue concept detection is depicted in Fig. 1. As mentioned in Sect. 1, one characteristic of our model is that it exploits different characteristics of words and tags, and offers tailor-made modeling for each of them. At the same time, the modeling of words and the modeling of tags are tied to each other in a coordinated manner. Precisely, an abstract venue concept is modeled as a probability distribution of tags, denoted by θ and a probability distribution of words, denoted by ϕ. Tags and words may have different vocabularies. The variable $|Z|$ denotes the number of abstract venue concepts.

Let D denote the set of venue profile documents. The outermost big plate in our graphical model represents a venue profile document, which contains a set of words, denoted by W and a set of tags, denoted by T. The number of words and tags is denoted by $|W|$ and $|T|$ respectively. Each venue profile

document is associated with a distribution of abstract concepts, denoted as φ. φ is assumed to be drawn from a Dirichlet distribution with a hyper-parameter α. φ has two components, namely tag concept assignment denoted as z_t, and word concept assignment denoted as z_w. Each tag t is associated with z_t. θ captures the distribution of tags for the concept represented by z_t.

Words in user comments are modeled in a different manner due to the different characteristics of user comments compared with tags. It is common that user comments may contain some unrelated content which has no relationship with the abstract venue concept at all. We employ a background distribution to model the general words in user comments, denoted by the variable b, which shares some resemblances with the modeling paradigm in [10]. The dimension of the variable b is the total number of words in the word vocabulary. User comments are treated as mixture of words in the background and words related to the abstract venue concept. Thus each word w is associated with either z_w or b, which is governed by a binary variable x. x is associated with a Bernoulli distribution λ with parameter ρ. The generative process of our model can be written as:

1. Draw φ from **Dirichlet**(α) and λ from **Beta**(ρ)
2. For each abstract venue concept
   i. Draw global tag distribution θ from **Dirichlet**(ς)
   ii. Draw global word distribution ϕ from **Dirichlet**(β)
3. For each venue profile document
   i. For each word $w \in W$ in the aggregated user comment
      a. Draw the word concept assignment z_w from **Multinomial**(φ)
      b. Draw switch x from **Bernoulli**(λ)
      c. Draw w from ϕ_{z_w} if $x = 1$, otherwise draw from b
   ii. For each tag $t \in T$
      a. Draw the tag concept assignment variable z_t from **Multinomial**(φ)
      b. Draw the tag t from θ_{z_t}

We use Gibbs sampling to compute the approximate posterior in our model. Let $|R_w|$ denote the number of tokens in the vocabulary built from user comments. Let $|R_t|$ denote the number of tokens in the vocabulary built from venue tags. Let $|B|$ denote the number of tokens in the background corpus. Let β_w denote an element in the hyper-parameter vector related to the word w. Let $n_{z_w w}$ denote number of times a word w in the user comment has been sampled from the abstract venue concept z_w. Similarly, let ς_t denote an element in a vector ς. Let $n_{z_t t}$ denote number of times a tag t in the tag vocabulary has been sampled from the abstract venue concept z_t. Let α_z represent accessing an element in the hyper-parameter vector α. Let q_{z_t} and q_{z_w} denote the number of times a global abstract venue concept has been sampled in a venue profile document. Note that when we have excluded the counts of the current case in our sampling equations. Let $\Theta = \{\boldsymbol{w}, \boldsymbol{t}, \beta, \alpha, \rho, \varsigma\}$. The complete likelihood of the model is denoted in Eq. 1.

$$P(\boldsymbol{z_w}, \boldsymbol{z_t}, \boldsymbol{w}, \boldsymbol{x}, \boldsymbol{t}|\alpha, \beta, \varsigma, \rho) = \int P(\boldsymbol{z_w}|\varphi)P(\varphi|\alpha)d\varphi \cdot \int P(\boldsymbol{z_t}|\varphi)P(\varphi|\alpha)d\varphi \cdot \int P(\boldsymbol{x}|\lambda)P(\lambda|\rho)d\lambda \cdot P(\boldsymbol{w}|\boldsymbol{x}, \boldsymbol{z_w}, \beta) \cdot \int P(\boldsymbol{t}|\boldsymbol{z_t}, \theta)P(\theta|\varsigma)d\theta \quad (1)$$

Where

$$P(\boldsymbol{w}|\boldsymbol{x} = 1, \boldsymbol{z_w}, \beta) = \int P(\boldsymbol{w}|\boldsymbol{z_w}, \phi)P(\phi|\beta)d\phi \quad (2)$$

$$P(\boldsymbol{w}|\boldsymbol{x} = 0, \boldsymbol{z_w}, \beta) = b_w \quad (3)$$

Eqs. 4, 5 and 6 depict the formulations used in our Gibbs sampler.

$$P(\boldsymbol{z_t}|\Theta) \propto \frac{\alpha_{z_t} + q_{z_t}}{\sum_{k=1}^{|Z|}(\alpha_k + q_k) - 1} \cdot \frac{\varsigma_t + n_{z_t t}}{\sum_{r=1}^{|R_t|}(\varsigma_r + n_{z_t r})} \quad (4)$$

$$P(\boldsymbol{z_w}|\boldsymbol{x} = 1, \Theta) \propto \frac{\alpha_{z_w} + q_{z_w}}{\sum_{k=1}^{|Z|}(\alpha_k + q_k) - 1} \cdot \frac{\beta_w + n_{z_w w}}{\sum_{r=1}^{|R_w|}(\beta_r + n_{z_w r})} \quad (5)$$

$$P(\boldsymbol{z_w}|\boldsymbol{x} = 0, \Theta) = \frac{b_w}{\sum_{k=1}^{|B|} b_k} \quad (6)$$

After sampling sufficient number of times, the parameters θ and ϕ are calculated with Eqs. 7 and 8.

$$\theta = \frac{\varsigma_t + n_{z_t t}}{\sum_{r=1}^{|R_t|}(\varsigma_r + n_{z_t r})} \quad (7) \qquad \phi = \frac{\beta_w + n_{z_w w}}{\sum_{r=1}^{|R_w|}(\beta_r + n_{z_w r})} \quad (8)$$

3 Labeling Abstract Venue Concepts

After the abstract venue concepts are detected, the next component is to automatically select one label to semantically describe the meaning of each concept.

For a discovered concept, the output from our model are a ranked list of tags and a ranked list of words from comments, obtained from matrices θ and ϕ respectively. The terms in the list coherently describe one venue concept. However, due to the intrinsic difference of tags and comments as shown in Sect. 1, these two lists generally contain some similar as well as different terms. For example, consider an abstract venue concept representing colleges, the corresponding abstract venue concept distribution for tags may consist of terms such as "library", "electronics", "college", "bookstore". Whereas the distribution of words in the abstract venue concept from the user comments may consist of "nice", "library", "excellent", "awesome", etc. Our objective is to automatically select representative tokens, such as "college" to serve as labels for that concept.

We adopt a technique based on the average Pointwise Mutual Information (PMI) described in [11], which also uses the same technique for finding topic labels. The value of PMI for a pair of words w_i and w_j is calculated with Eq. 9, where $P(w_i, w_j)$ denotes the probability of observing both w_i and w_j in the same list. $P(w_i)$ and $P(w_j)$ are the overall probability of token w_i and w_j respectively. Then the average PMI is calculated by averaging over all the tokens in the list, denoted by Eq. 10. PMI measures the association between one event to other events using information theory and statistics. In our case, intuitively, tokens that has more co-occurrence with other tokens will get higher PMI. For each discovered concept, we basically choose two concept labels that have the highest average PMI from the tag list and word list. We select the word with the highest avgPMI from the two ranked lists discovered by our model.

$$PMI = \log \frac{P(w_i, w_j)}{P(w_i) \cdot P(w_j)} \tag{9}$$

$$avgPMI = \frac{1}{N} \sum_{j}^{N} PMI(w_i, w_j) \tag{10}$$

4 Experiments and Results

4.1 Datasets

We used the official Foursquare API to crawl text data related to tags and user comments corresponding to that venue. We crawled data from the following countries and in brackets we list the number of venue profile documents: (1) Australia (30,880), (2) Canada (50,063), (3) Hong Kong (5,282), (4) India (12,277), (5) Indonesia (302,725), (6) Singapore (18,082), and (7) USA (879,476). We selected those venues which had text content in both tags and user comments. Each venue document obtained from Foursquare contains several tags and up to 20 comments.

4.2 Comparative Models

We choose a range of comparative models including some state-of-the-art topic models. Specifically, we compare our proposed model denoted as "Our Model" with (1) Latent Dirichlet Allocation (`LDA`) model [9]. We compare with both variational inference [9] and collapsed Gibbs sampling based algorithms [12] denoted by `vLDA` and `cLDA` respectively. (2) Topical N-gram (`TNG`) [13,14] model which is a phrase discovery topic model, (3) Hierarchical Dirichlet Processes (`HDP`) topic model [15,16], which is a nonparametric extension to the `LDA` model, (4) Biterm topic model (`BTM`) [17,18], which is a topic model suited for short texts as most of the documents in our collection are short. We use publicly available source codes of all these models. We used the same parameter settings of these models as described in their respective works. We use fixed symmetric Dirichlet distributions in our model in which we set $\alpha = 0.5$, $\beta = 0.01$, $\varsigma = 0.01$. In addition, we

fixed $\rho = 0.01$ in our model. All models are run for 1000 iterations. We combine user comments and tags in one document for the comparative methods as used in [8].

4.3 Concept Coherence Evaluation

The first evaluation measures the quality of concepts generated by the models. To enable large scale evaluation, we evaluate topical coherence using an automated technique called observed coherence model discussed in [19]. The idea is to automatically find out whether the list of tokens in each concept are semantically related, which in turn leads to better concept interpretability.

In all models, we varied the number of concepts from 10 to 200 in steps of 10 except the HDP model which automatically finds out the number of latent concepts. We run the topic models for five times due to randomization as adopted in [20]. Therefore, for each concept, each model was run for five times and the average coherence score was computed in each run. Then the macro-average coherence score was computed for all five runs. We then computed the average across different number of concepts from 10 to 200.

Table 2. Average coherence scores obtained for different models in different datasets. The higher the average coherence score, the better is the model.

	vLDA	cLDA	TNG	HDP	BTM	Our Model
Australia	0.120	0.150	0.110	0.090	0.002	**0.220**
Canada	0.020	0.040	0.006	0.001	0.005	**0.120**
Hong Kong	0.013	0.300	0.120	0.090	0.002	**0.210**
India	0.010	0.020	0.090	0.002	0.006	**0.200**
Indonesia	0.011	0.013	0.008	0.008	0.003	**0.320**
Singapore	0.009	0.020	0.020	0.005	0.003	**0.140**
USA	0.003	0.040	0.010	0.009	0.003	**0.110**

We present the results obtained from different models in Table 2. We see from the table that our model has obtained the best average coherence score with improvements that is statistically significant according to two-tailed test with $p < 0.01$ against each of the comparative models. One may argue that comparative models may perform better if we separately model user comments and venue tags as separate documents. We found that results obtained from such strategy are even worse due to the sparsity problem. Our model jointly models words from user comments along with other useful information from venues, leading to more coherent concepts which mitigates the sparsity issue. In addition, introducing the background distribution helps us get rid of many irrelevant words which were dominant in many comparative models. Presenting only the average performance for all topics hides the per-topic performance of a model, but it must be noted that our model performs consistently better than the comparative models at different number of topics.

4.4 Concept Label Evaluation

We also evaluate the quality of the concept labeling task. We hired five human annotators to give ratings to concept labels. In our annotation task, each concept was presented in the form of its top-20 words ranked with decreasing probability value, followed by suggested label for the concept. Since our model generates two word distributions each of which will have a label of its own. In order to reduce the cognitive load on the human annotators, we only gave them the list of five concepts from each model, i.e. $|Z| = 5$. For HDP, five concepts were randomly selected. We gave the ordinal scale rating questions to the annotators as described in [21]. The ratings range from 0 to 3. We considered the scores as voting given by the human annotators and computed the average score from all annotators for each model.

Table 3. Average ratings given by annotators. The higher the average score, the better is the model.

	vLDA	cLDA	TNG	HDP	BTM	Our Model
Australia	1.56	1.32	1.12	0.28	0.45	**2.80**
Canada	1.68	1.04	0.96	0.40	0.20	**2.60**
Hong Kong	1.04	1.20	1.04	0.40	0.20	**2.40**
India	1.16	0.76	1.36	0.16	0.30	**2.40**
Indonesia	1.40	1.04	0.72	1.08	0.42	**2.15**
Singapore	0.92	0.80	1.20	0.28	0.16	**2.75**
USA	0.76	1.00	0.92	0.60	0.22	**2.40**

We present the results in Table 3. We see from the results that our model has obtained the highest value compared with other models.

Specially, the standard LDA methods, i.e. vLDA and cLDA, which aggregate tags and comments and do not model them differently, perform worse than our model. This observation shows the advantage of our tailor-made modeling, which exploits different characteristics of tags and comments. In addition, the relatively worse performances of HDP, compared with LDA, show that the venue profile documents do not have distinct hierarchical properties. The BTM, which is suitable for short texts, fails to get better results, although most of the documents in our collection are short. TNG shows comparable performance with standard LDA.

4.5 Sample Concept Case Study

We present top 20 terms from some abstract venue concepts from our model discovered from the "Australia" dataset in Table 4. We have merged the top ten words from two lists output by our model, and arranged the words in decreasing order of their probability values in the list presented below. There are two terms,

Table 4. Top 20 terms for some concepts. Labels are in Bold font

Top 20 terms and labels
asian, **restaurant**, indonesian, noodles, ramen, chinese, seafood, pizza, food,caf, diner, japanese, arcade, bbq, soup, breakfast, steakhouse, sushi, italian
university, video, games, music, **library**, electronics, building, school, store, **college**, office, education, bookstore, books, academic, tv, dvd, bowling, camera
theater, movie, theatre, apparel, cineplex, music, arts, movies, concert, performing, bowling,**entertainment**, hall, art, winery, popcorn, venue, gallery, alley
park, **playground**, **outdoors**, golf, hotel, field, beach, baseball, dog pool, lake,trail, apartments, museum, run, boat, scenic, entertainment, lookout

one from each distribution, selected as labels which are in bold font. We see from the sample terms that our model has generated meaningful and coherent terms. For example, the second concept, which is labeled with "college" and "library", apparently represents a college. Most of the words are related to college where university students play video games in the residential buildings. They go to the library to read books, listen to music, watch television or DVD, etc.

Our model has generated more superior results than the comparative models because of the following reasons. First, the background distribution in our model helps get rid of many general words from the distributions. This helps focus on only relevant content words in the user comments. Comparative models such as `LDA`, `TNG`, etc. are not designed to handle this. Although we could adopt aggressive pre-processing methods and then input the pre-processed text to the comparative models, it involves manual labour to select such general terms and removing them. Automatically removing the general words from the corpus can also be adopted, for example, using term-frequency and inverse document frequency score and removing those words which have low scores. But this involves selecting an appropriate threshold value, and we need to expend some computing time too. We could have considered a background distribution in the comparative models by modifying the models slightly and their sampling algorithms, however, those models still lack the ability to separately model user comments and tags in order to generate high quality abstract venue concepts.

5 Conclusion

We have proposed a new model to generate abstract venue concepts from LBSNs venue profiles. Our model jointly models user comment text and tags. Meanwhile, the model offers tailor-made modeling for these two kinds of text data, exploiting their different characteristics.We conducted extensive experiments and found our model to be superior compared with comparative models in both the coherence of concept and the quality of labels.

References

1. Yuan, Q., Cong, G., Sun, A.: Graph-based point-of-interest recommendation with geographical and temporal influences. In: Proceedings of the 23rd ACM International Conference on Information and Knowledge Management, pp. 659–668 (2014)
2. Zhong, Y., Yuan, N.J., Zhong, W., Zhang, F., Xie, X.: You are where you go: inferring demographic attributes from location check-ins. In: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, pp. 295–304 (2015)
3. Yuan, N.J., Zhang, F., Lian, D., Zheng, K., Yu, S., Xie, X.: We know how you live: exploring the spectrum of urban lifestyles. In: Proceedings of the First ACM Conference on Online Social Networks, pp. 3–14 (2013)
4. Liu, B., Xiong, H.: Point-of-interest recommendation in location based social networks with topic and location awareness. In: Proceedings of the 2013 SIAM International Conference on Data Mining, pp. 396–404 (2013)
5. Wang, C., Wang, J., Xie, X., Ma, W.Y.: Mining geographic knowledge using location aware topic model. In: Proceedings of the 4th ACM Workshop on Geographical Information Retrieval, pp. 65–70 (2007)
6. Wang, X., Zhao, Y.L., Nie, L., Gao, Y., Nie, W., Zha, Z.J., Chua, T.S.: Semantic-based location recommendation with multimodal venue semantics. IEEE Trans. Multimedia **17**(3), 409–419 (2015)
7. Hong, L., Ahmed, A., Gurumurthy, S., Smola, A.J., Tsioutsiouliklis, K.: Discovering geographical topics in the twitter stream. In: Proceedings of the 21st International Conference on World Wide Web, pp. 769–778 (2012)
8. Kim, E., Ihm, H., Myaeng, S.H.: Topic-based place semantics discovered from microblogging text messages. In: Proceedings of the Companion Publication of the 23rd International Conference on World Wide Web Companion, pp. 561–562 (2014)
9. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (2003)
10. Chemudugunta, C., Steyvers, P.S.M.: Modeling general and specific aspects of documents with a probabilistic topic model. In: Proceedings of the 2006 Conference in Neural Information Processing Systems 19, vol. 19, p. 241 (2007)
11. Lau, J.H., Newman, D., Karimi, S., Baldwin, T.: Best topic word selection for topic labelling. In: Proceedings of the 23rd International Conference on Computational Linguistics: Posters, pp. 605–613 (2010)
12. Porteous, I., Newman, D., Ihler, A., Asuncion, A., Smyth, P., Welling, M.: Fast collapsed gibbs sampling for latent Dirichlet allocation. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 569–577 (2008)
13. Wang, X., McCallum, A., Wei, X.: Topical n-grams: Phrase and topic discovery, with an application to information retrieval. In: Seventh IEEE International Conference on Data Mining, pp. 697–702. IEEE (2007)
14. Wang, X., McCallum, A.: A note on topical n-grams. Technical report, DTIC Document (2005)
15. Teh, Y.W., Jordan, M.I., Beal, M.J., Blei, D.M.: Hierarchical Dirichlet processes. J. Am. Stat. Assoc. **101**(476), 1566–1581 (2006)
16. Teh, Y.W., Kurihara, K., Welling, M.: Collapsed variational inference for HDP. In: Advances in Neural Information Processing Systems, pp. 1481–1488 (2007)

17. Yan, X., Guo, J., Lan, Y., Cheng, X.: A biterm topic model for short texts. In: Proceedings of the 22nd International Conference on World Wide Web, pp. 1445–1456 (2013)
18. Cheng, X., Yan, X., Lan, Y., Guo, J.: BTM: topic modeling over short texts. IEEE Trans. Knowl. Data Eng. **26**(12), 2928–2941 (2014)
19. Lau, J.H., Newman, D., Baldwin, T.: Machine reading tea leaves: automatically evaluating topic coherence and topic model quality. In: Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, p. 530 (2014)
20. Zhu, J., Ahmed, A., Xing, E.P.: MedLDA: maximum margin supervised topic models. J. Mach. Learn. Res. **13**(1), 2237–2278 (2012)
21. Lau, J.H., Grieser, K., Newman, D., Baldwin, T.: Automatic labelling of topic models. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, vol. 1, pp. 1536–1545 (2011)