



A survey on neural topic models: methods, applications, and challenges

Xiaobao Wu¹ · Thong Nguyen² · Anh Tuan Luu¹

Accepted: 19 December 2023 / Published online: 25 January 2024
© The Author(s) 2024

Abstract

Topic models have been prevalent for decades to discover latent topics and infer topic proportions of documents in an unsupervised fashion. They have been widely used in various applications like text analysis and context recommendation. Recently, the rise of neural networks has facilitated the emergence of a new research field—neural topic models (NTMs). Different from conventional topic models, NTMs directly optimize parameters without requiring model-specific derivations. This endows NTMs with better scalability and flexibility, resulting in significant research attention and plentiful new methods and applications. In this paper, we present a comprehensive survey on neural topic models concerning methods, applications, and challenges. Specifically, we systematically organize current NTM methods according to their network structures and introduce the NTMs for various scenarios like short texts and cross-lingual documents. We also discuss a wide range of popular applications built on NTMs. Finally, we highlight the challenges confronted by NTMs to inspire future research.

Keywords Neural topic model · Cross-lingual topic modeling · Dynamic topic modeling · Short text topic modeling · Variational AutoEncoder

1 Introduction

Topic models seek to discover a set of latent topics from a collection of documents, depending on word co-occurrence information. Each topic represents an interpretable semantic concept and is described as a group of related words. For example, a topic about

✉ Xiaobao Wu
xiaobao002@e.ntu.edu.sg

✉ Anh Tuan Luu
anhtuan.luu@ntu.edu.sg

Thong Nguyen
e0998147@u.nus.edu

¹ School of Computer Science and Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore

² School of Computing, National University of Singapore, 21 Lower Kent Ridge Rd, Singapore 119077, Singapore

“sports” may relate to words like “baseball”, “basketball”, and “football”. Topic models also infer what topics a document contains to reveal their underlying semantics. Due to their effectiveness and interpretability, topic models have derived various downstream applications, such as document retrieval, content recommendation, opinion/event mining, and trend analysis (Blei and Lafferty 2006b; Wang and Blei 2011; Boyd-Graber et al 2017; Duong et al 2022; Churchill and Singh 2022).

Conventional approaches to topic modeling embrace either probabilistic graphical models or non-negative matrix factorization. Approaches based on probabilistic graphical models, such as Latent Dirichlet Allocation (LDA, Blei et al 2003), have been extensively explored for the past two decades. They mainly model the document generation process with topics as latent variables (Blei 2012). Then they infer model parameters through Variational Inference (Blei et al 2017) or Monte Carlo Markov Chain (MCMC) methods like Gibbs sampling (Steyvers and Griffiths 2007). Alternatively, another conventional type of topic models uses non-negative matrix factorization. They directly discover topics by decomposing a term-document matrix into two low-rank factor matrices: one represents words and the other documents (Lee and Seung 2000; Kim et al 2015; Shi et al 2018). These conventional topic models have derived various model structures, such as supervised LDA (Mcauliffe and Blei 2007) and correlated LDA (Blei and Lafferty 2006a). Besides the basic topic modeling scenario, researchers have extended topic models to other diverse scenarios, e.g., short text (Yan et al 2013; Yin and Wang 2014), cross-lingual (Mimno et al 2009), and dynamic topic modeling (Blei and Lafferty 2006b; Wang et al 2008).

However, despite the achievements of these conventional methods, they generally confront two limitations: (i) **Inefficient and labor-intensive parameter inference**. These methods necessitate complicated model-specific derivations for parameter inference, and the inference complexity grows along with model complexity. Consequently, this requirement weakens their generalization ability to diverse model structures and application scenarios. (ii) **Limited scalability to large datasets**. Their inference algorithms typically are *not* parallel, leading to significant time consumption. For example, training a probabilistic dynamic topic model on a dataset with 10k documents takes two days (Dieng et al 2019). Admittedly some parallel inference algorithms have been proposed (Newman et al 2009; Wang et al 2009; Liu et al 2011), but unfortunately they cannot straightforwardly fit other model structures and application scenarios. As a result, how to design effective, flexible, efficient, and scalable topic models has become an urgent imperative.

To overcome these challenges, Neural Topic Models (NTMs) have emerged as a promising solution. Unlike conventional topic models, NTMs can efficiently and flexibly infer the model parameters through automatic gradient back-propagation by adopting deep neural networks to model latent topics, such as the popular Variational AutoEncoder (VAE, Kingma and Welling 2014; Rezende et al 2014). This flexibility enables researchers to tailor model structures to fit diverse application scenarios. In addition, NTMs can seamlessly handle large-scale datasets by harnessing parallel computing facilities like GPUs. Owing to these advantages, NTMs have witnessed the exploration of numerous new methods and applications.

Previously, Zhao et al (2021a) provided a review with a primary focus on the methods of NTMs. However, their review is beset by the following limitations: (i) Their method taxonomy is incomplete because they ignore several recently proposed NTM methods, such as NTMs with contrastive learning, cross-lingual NTMs, and dynamic NTMs. (ii) They omit the popular applications based on NTMs, developed for a wide range of downstream tasks. (iii) They lack in-depth discussions on the challenges inherent in NTMs. As a consequence, a more comprehensive review on NTMs is necessary for the research field.

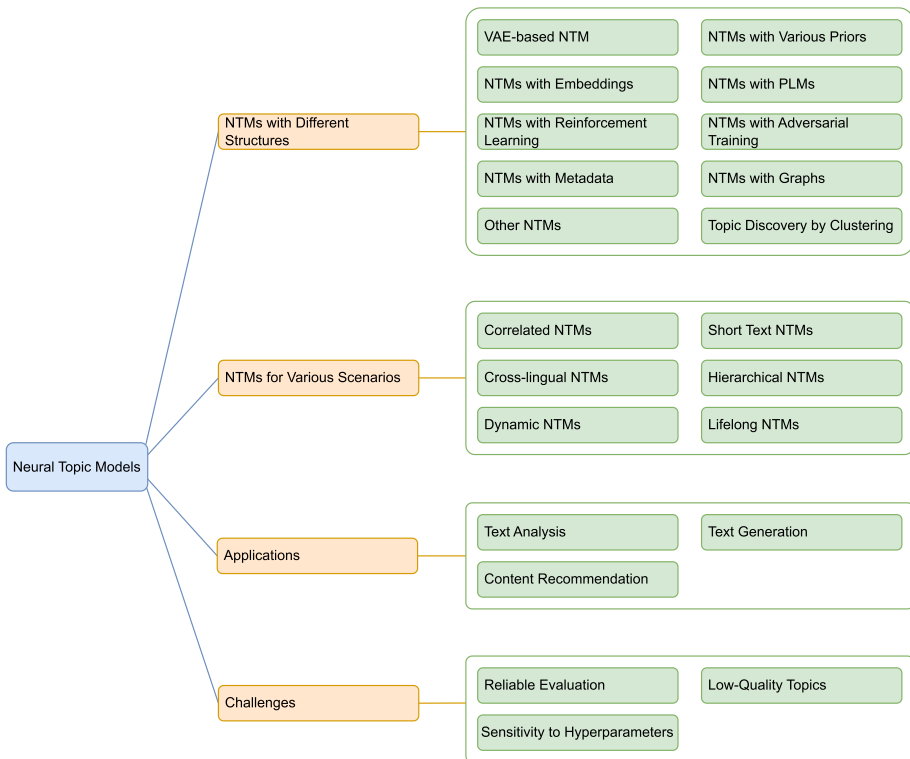


Fig. 1 The overview of this survey, including NTMs with different structures, NTMs for various scenarios, applications, and challenges

To address these limitations, we in this paper present an extensive and up-to-date survey of NTMs, which offers an in-depth and self-contained understanding of NTMs in terms of methods, applications, and challenges. We begin by systematically organizing existing NTMs according to their neural network structures, such as using embeddings or graph neural networks. We then introduce the NTMs designed for various prevalent topic modeling scenarios, e.g., short text, cross-lingual, and dynamic topic modeling, covering a wider range than the early survey (Zhao et al 2021a). Moreover while omitted by the previous survey (Zhao et al 2021a), we also organize and discuss the popular applications based on NTMs, developed for diverse tasks like text analysis and text generation. Finally we summarize the key research challenges for NTMs in detail to motivate future research directions. Fig. 1 depicts the overview of our survey. We conclude the main contributions of this paper as follows:

- We extensively review methods of neural topic models through detailed discussions and comparisons, covering variants with different network structures.
- We include a broader range of popular topic modeling scenarios and provide detailed background information for each scenario, accompanied by easy-to-understand illustrations and related neural topic models.

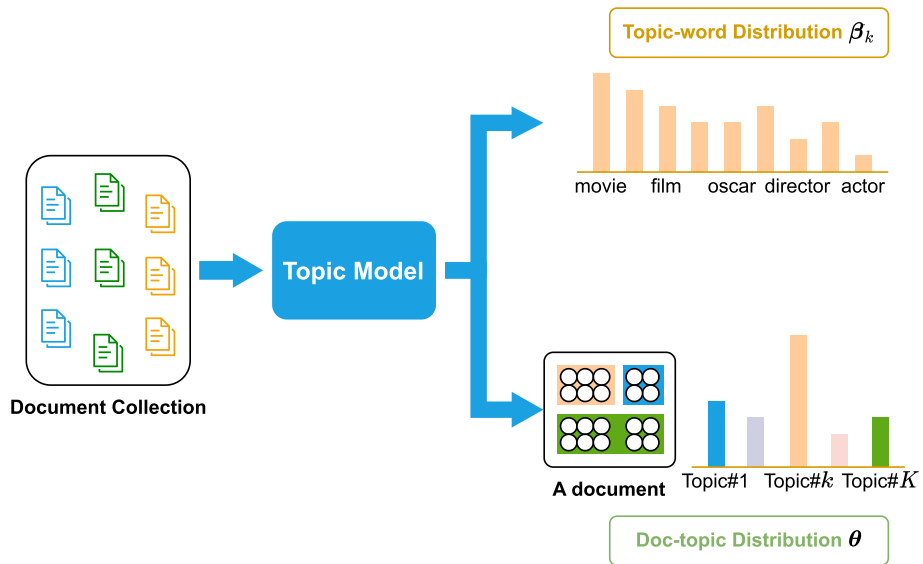


Fig. 2 Illustration of topic modeling. Given a document collection, a topic model aims to discover K latent topics, interpreted as distributions over words (topic-word distributions). It also infers what topics a document contains, defined as distributions over all latent topics (doc-topic distributions). Here the topic-word distribution of Topic# k , β_k , has related words like “movie”, “film”, and “oscar”; the doc-topic distribution θ concentrates on Topic#1 and Topic# k

- We introduce popular applications based on neural topic models, developed to tackle various tasks such as text analysis and generation.
- We highlight the current vital challenges faced by neural topic models in detail to facilitate future research; Motivated by this, we propose a new topic diversity metric that measures diversity along with word semantics, which more agrees with human judgment.

We accompany this survey with a repository ¹ of the mentioned paper resources to provide easy access for researchers.

2 Preliminary

In this section, we introduce the preliminary of topic modeling, including the problem setting, notations, and evaluation methods. Then we present the most basic and popular NTM in the framework of Variational AutoEncoder (VAE).

¹ <https://github.com/bobxwu/Paper-Neural-Topic-Models>.

2.1 Problem setting and notations

We introduce the problem setting and notations of topic modeling following LDA (Blei et al 2003). Consider a collection of N documents with V unique words (vocabulary size), and a document is denoted as \mathbf{x} . As illustrated in Fig. 2, topic models aim to discover K latent topics from this collection. The number of topics K is a hyperparameter, usually determined by researchers manually according to the characteristics of datasets and their target tasks. Each topic is defined as a distribution over the vocabulary, *i.e.*, **topic-word distribution**, $\beta_k \in \mathbb{R}^V$. Then we have $\beta = (\beta_1, \dots, \beta_K) \in \mathbb{R}^{V \times K}$ as the topic-word distribution matrix of all topics. In addition, topic models also infer the topic distribution of a document (**doc-topic distribution**): $\theta \in \Delta_K$, implying what topics a document contains. Here θ_k refers to the proportion of Topic# k in the document, and Δ_K denotes a probability simplex $\Delta_K = \{\theta \in \mathbb{R}_+^K \mid \sum_{k=1}^K \theta_k = 1\}$.

2.2 Evaluation of topic models

Given the absence of ground-truth labels in topic modeling tasks, how to reliably and comprehensively evaluate topic models remains inconclusive in the research community. We introduce currently the most prevalent evaluation methods employed for assessing topic models as follows.

2.2.1 Perplexity

Perplexity, borrowed from language models, measures how a model can predict new documents. It is measured as the normalized log-likelihood of held-out test documents. Perplexity has been used for years to evaluate topic models. Nevertheless, prior studies have empirically demonstrated that perplexity inaccurately reflects the quality of discovered topics as it often contradicts human judgment (Chang et al 2009). Furthermore, computing log-likelihood is inconsistent among different topic models. This is because they apply various sampling or approximation techniques (Wallach et al 2009; Buntine 2009) as well as diverse modeling approaches for topic-word distributions and doc-topic distributions. For instance, certain methods normalize topic-word distributions with respect to topics, some with respect to words, and others opt to keep them unnormalized. These disparities bring challenges to equitable comparisons. Finally, perplexity may not evaluate the practical utility of topic models since users typically employ topic models for content analysis rather than generating new documents (Zhao et al 2021a; Hoyle et al 2022). Due to these reasons, perplexity has waned in popularity for topic model evaluation in the recent research field.

2.2.2 Topic coherence

Rather than predictive abilities, researchers turn to evaluating the quality of produced topics. For this purpose, researchers propose topic coherence to measure the coherence among the most related words of topics, *i.e.*, top words (determined by topic-word probabilities). Experiments showcase that topic coherence can agree with the human evaluation on topic interpretability (Lau et al 2014). For example, one widely-used coherence metric is Normalized Point-wise Mutual Information (NPMI, Bouma 2009; Newman et al 2010;

Lau et al (2014).² Specifically, the NPMI score between two words (x_i, x_j) is calculated as follows:

$$\text{NPMI}(x_i, x_j) = \frac{\log \frac{p(x_i, x_j) + \epsilon}{p(x_i)p(x_j)}}{-\log p(x_i, x_j) + \epsilon}. \quad (1)$$

It computes the normalized mutual information of two words, and then takes the average of all word pairs in all topics. Here ϵ is to avoid zero; $p(x_i)$ is the probability of word x_i , and $p(x_i, x_j)$ is the co-occurrence probability of (x_i, x_j) . These probabilities are estimated as their occurrence frequencies in a reference corpus. The reference corpus can be either internal (the training set) or external (e.g., Wikipedia articles). Basically, a large external corpus is recommended because it can alleviate the influence of data bias in training sets and facilitate fair topic coherence comparisons across different datasets.

Later, Röder et al (2015) propose a new metric, C_V , which calculates the cosine similarity between NPMI score vectors (Krasnashchok and Jouili 2018). Given the top T words of a topic, (x_1, x_2, \dots, x_T) , the exact calculation of C_V is formulated as

$$C_V = \frac{1}{T} \sum_{i=1}^T \cos(\mathbf{v}_{\text{NPMI}}(x_i), \mathbf{v}_{\text{NPMI}}(\{x_i\}_{i=1}^T)) \quad (2)$$

$$\mathbf{v}_{\text{NPMI}}(x_i) = \{\text{NPMI}(x_i, x_j)\}_{j=1, \dots, T} \quad (3)$$

$$\mathbf{v}_{\text{NPMI}}(\{x_i\}_{i=1}^T) = \left\{ \sum_{i=1}^T \text{NPMI}(x_i, x_j) \right\}_{j=1, \dots, T}. \quad (4)$$

The NPMI score computation follows Equation (1). Röder et al (2015) empirically demonstrate that C_V outperforms previous coherence metrics, NPMI, UCI, and UMass (Mimno et al 2011), since C_V more agrees with human judgment (See Röder et al (2015) for experimental results).

We would like to recommend the Palmetto tool³ to compute topic coherence. It includes almost all common coherence metrics and provides a pre-processed Wikipedia article collection as the reference corpus for easier reproducibility.

2.2.3 Topic diversity

To further evaluate the quality of topics, topic diversity is introduced to measure the difference between topics. This is driven by the anticipation that topics should exhibit diversity rather than redundancy thereby enabling the comprehensive disclosure of latent semantics in corpora. At present, researchers employ the following diversity metrics:

² https://github.com/jhlau/topic_interpretability.

³ <https://github.com/dice-group/Palmetto>.

- Nan et al (2019) propose Topic Uniqueness (TU) which computes the average reciprocal of top word occurrences in topics. In detail given K topics and the top T words of each topic, TU is computed as

$$\text{TU} = \frac{1}{K} \sum_{k=1}^K \frac{1}{T} \sum_{x_i \in t(k)} \frac{1}{\#(x_i)} \quad (5)$$

where $t(k)$ means the top word set of the k -th topic, and $\#(x_i)$ denotes the occurrence of word x_i in the top T words of all topics. TU ranges from $1/K$ to 1.0, and a higher TU score indicates more diverse topics.

- Burkhardt and Kramer (2019) propose Topic Redundancy (TR) that calculates the average occurrences of a top word in other topics. Its computation is

$$\text{TR} = \frac{1}{K} \sum_{k=1}^K \frac{1}{T} \sum_{x_i \in t(k)} \frac{\#(x_i) - 1}{K - 1}. \quad (6)$$

A higher TR score means less diverse topics.

- Dieng et al (2020) propose Topic Diversity (TD) which computes the proportion of unique top words of topics:

$$\text{TD} = \frac{1}{K} \sum_{k=1}^K \frac{1}{T} \sum_{x_i \in t(k)} \mathbb{1}(\#(x_i)) \quad (7)$$

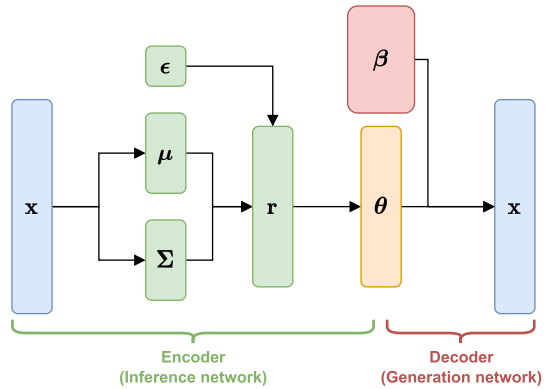
where $\mathbb{1}(\cdot)$ is a indicator function that equals 1 if $\#(x_i) = 1$ and equals 0 otherwise. TD ranges from 0 to 1.0, and a higher TD score indicates more diverse topics.

These metrics all measure topic diversity based on the uniqueness of individual words. They posit that diversity is optimal when all topics are characterized by distinct top words. However, we question these diversity metrics because of the fact that certain topics naturally share the same words. For example, the word “chip” could be shared by the topics of “potato chip” and “electronic chip”; Similarly, the word “apple” may be covered by the topics of “fruit” and “company”. This issue remains unresolved for reliable diversity evaluation. We in this paper propose a new diversity metric to address this issue (See details in Sect. 6).

2.2.4 Downstream task performance

Except for the coherence and diversity to measure topic quality, researchers also resort to extrinsic performance: they use doc-topic distributions θ as low-dimensional document features and evaluate their quality on downstream tasks. These tasks mainly consist of document classification and document clustering. For document classification, researchers train ordinary classifiers (e.g., SVMs or Random Forests) with learned doc-topic distributions as document features and then predict the labels of testing documents. The performance can be evaluated by accuracy or F1 scores. For document clustering, the common way is to use the most significant topic in a doc-topic distribution as the clustering assignment of a document. Another way is to apply clustering algorithms, e.g., K-Means or DBSCAN, on doc-topic distributions (Zhao et al 2021b). The clustering performance can be measured by Purity and Normalized Mutual Information (NMI, Manning et al 2008).

Fig. 3 Illustration of a VAE-based NTM. It mainly contains an encoder (inference network) and a decoder (generation network). The encoder outputs doc-topic distribution θ from input document \mathbf{x} through MLPs using the reparameterization trick where $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. The decoder reconstructs the input document from θ with β as the topic-word distribution matrix. The objective includes reconstruction error and KL divergence



2.2.5 Visualization

Finally, researchers visualize topic models for evaluation. The typical visualization method is to show the top words of topics and doc-topic distributions, such as using pyLDAvis⁴ (Sievert and Shirley 2014) or word cloud.⁵ Another strategy is to cluster documents on a 2D canvas by reducing the dimension of doc-topic distributions with tools like t-SNE (van der Maaten and Hinton 2008).

2.3 Basic NTM based on VAE

We introduce the most basic and popular NTM based on the Variational AutoEncoder (VAE) framework with the neural variational inference technique (Miao et al 2016; Srivastava and Sutton 2017). As illustrated in Fig. 3, a VAE-based NTM mainly contains an encoder (inference network) and a decoder (generation network). The encoder is to infer doc-topic distributions from documents. To be specific, we use a latent variable $\mathbf{r} \in \mathbb{R}^K$ following a logistic normal prior

$$p(\mathbf{r}) = \mathcal{LN}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \tag{8}$$

where $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ are the mean vector and diagonal covariance matrix respectively. Here the prior distribution is specified with Laplace approximation (Hennig et al 2012) to approximate a symmetric Dirichlet prior as $\mu_{0,k} = 0$ and $\Sigma_{0,kk} = (K - 1)/\alpha K$ with hyperparameter α (Srivastava and Sutton 2017). The variational distribution is modeled as

$$q_{\Theta}(\mathbf{r}|\mathbf{x}) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}). \tag{9}$$

We compute $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ with encoder networks parameterized by Θ :

$$\boldsymbol{\mu} = f_{\Theta_1}(\mathbf{x}) \tag{10}$$

⁴ <https://github.com/bmabey/pyLDAvis>.

⁵ https://github.com/amueller/word_cloud.

$$\Sigma = \text{diag}(f_{\Theta_2}(\mathbf{x})) \tag{11}$$

where $\Theta = \{\Theta_1, \Theta_2\}$ and $\text{diag}(\cdot)$ denotes transforming a vector to a diagonal matrix. In practice, we transform document \mathbf{x} into a Bag-of-Words (BoW) vector as the input and employ MLPs as encoder networks. Then to avoid gradient variance (Kingma and Welling 2014; Rezende et al 2014), we sample \mathbf{r} through the reparameterization trick by sampling a random variable ϵ :

$$\mathbf{r} = \boldsymbol{\mu} + (\Sigma)^{1/2}\epsilon \quad \text{where} \quad \epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I}). \tag{12}$$

We model the doc-topic distribution θ with a softmax function to restrict it on a simplex:

$$\theta = \text{softmax}(\mathbf{r}). \tag{13}$$

The decoder is to generate documents from doc-topic distributions. Specifically, we use a decoder network parameterized by Φ : $f_{\Phi}(\theta) = \text{softmax}(\beta\theta)$ which represents the generation probability of each word. Here $\Phi = \{\beta\}$. Then we sample words from its multinomial distribution: $x \sim \text{Mult}(f_{\Phi}(\theta))$. Following the Evidence Lower Bound (ELBO) of VAE, we formulate the learning objective of NTM as

$$\min_{\Theta, \Phi} -\mathbb{E}_{q_{\Theta}(\theta|\mathbf{x})} [\log p_{\Phi}(\mathbf{x}|\theta)] + \text{KL}[q_{\Theta}(\mathbf{r}|\mathbf{x})\|p(\mathbf{r})]. \tag{14}$$

The first term is the negative expected log-likelihood, i.e., the reconstruction error, where $p_{\Phi}(\mathbf{x}|\theta)$ denotes the generation probability of \mathbf{x} . As we sample words from the multinomial distribution, the first term becomes $-\mathbf{x}^T \log(f_{\Phi}(\theta))$. The second term is the Kullback–Leibler (KL) divergence between the variational and prior distributions, which can be computed through an analytical form (Srivastava and Sutton 2017). It is also known as a regularization term.

The above is the fundamental structure of a VAE-based NTM. Based on this, NTMs with different structures are proposed to further improve performance and deal with various application scenarios.

3 Review of neural topic models

In this section, we review the existing work of Neural Topic Models (NTMs). We first introduce the NTMs with different structures, and then discuss the NTMs for various use case scenarios.

3.1 NTMs with different structures

Apart from the basic VAE structure mentioned in Sect. 2, we introduce NTMs with more different structures.

3.1.1 NTMs with various priors

VAE-based NTMs commonly employ Gaussian (Normal) as priors since it is easy to apply the reparameterization trick and compute the analytical form of KL divergence. Besides Gaussian priors, NTMs also leverage other various priors. Miao et al (2017) propose new

priors like Gaussian softmax and the stick-breaking process. Zhang et al (2018) use a Weibull distribution to approximate gamma distributions. Joo et al (2020) leverage an auxiliary uniform distribution to approximate the cumulative distribution function of gamma. As Dirichlet priors are important for topic modeling, Burkhardt and Kramer (2019) utilize the proposal function of a rejection sampler for a gamma distribution to approximate Dirichlet priors. Tian et al (2020) draw from the rounded posterior distribution to approximate Dirichlet samples.

3.1.2 NTMs with embeddings

Alternative to directly modeling topics, Miao et al (2017) propose to decompose topics as two embedding parameters:

$$\beta = \mathbf{W}^T \mathbf{T}. \quad (15)$$

Here $\mathbf{W} \in \mathbb{R}^{D \times V}$ denotes V word embeddings, and $\mathbf{T} \in \mathbb{R}^{D \times K}$ means K topic embeddings, where D is the dimension of embedding space. Then Dieng et al (2020) follow this setting and facilitate topic learning by initializing \mathbf{W} with pre-trained word embeddings like Word2Vec (Mikolov et al 2013) or GloVe (Pennington et al 2014). This approach also confers flexibility and efficiency to other topic modeling scenarios. For instance in dynamic topic modeling, it is much cheaper to repeat topic embeddings for each time slice than repeating the entire topic-word distribution matrix (Dieng et al 2019).

Alternatively, Zhao et al (2021b) also model topics as embeddings, but use the optimal transport distance between doc-topic distributions and input documents to measure the reconstruction error. Wang et al (2022a) share the same idea and instead use conditional transport distance. Duan et al (2022) learn a group of global topic embeddings for task-specific adaptations. Xu et al (2022) use hyperbolic embeddings to model topics. Due to the tree-likeness property of hyperbolic space, they can capture the hierarchy among topics. More recently, Wu et al (2023b) find that topic embeddings mostly collapse together in the space of NTMs with embeddings, which leads to topic collapsing, *i.e.*, repetitive topics. To address this issue, they propose a regularization on embeddings in addition to the traditional objective based on ELBO. The regularization considers topic embeddings as cluster centers and word embeddings as cluster samples; then it forces topic embeddings to be the centers of separately aggregated word embeddings. This effectively mitigates the topic collapsing issue and extensively improves topic modeling performance.

3.1.3 NTMs with metadata

While common NTMs learn topics in an unsupervised manner, NTMs can also leverage the metadata of documents to guide topic modeling, similar to supervised LDA (Mcauliffe and Blei 2007). In detail, Card et al (2018) introduce a NTM that can incorporate various metadata of documents. It encodes a document together with its labels (e.g., sentiment) and covariates (e.g., publication year), and generates the document conditioned on the covariates. Korshunova et al (2019) model the generation of documents and labels together in a discriminative way; then train their model with mean-field variational inference. They can also incorporate a variety of data modalities like images. Wang and Yang (2020) jointly model topics and train a RNN classifier to predict document labels. They are connected by an attention mechanism. Wang et al (2021a) incorporate document networks in a NTM and jointly reconstruct documents and networks.

3.1.4 NTMs with graph neural networks

In addition to traditional BoW as inputs, several NTMs use graph neural networks to model documents. Specifically, Zhu et al (2018) transform documents into biterm graphs and follow the VAE framework to reconstruct the input graphs. A biterm refers to an unordered word pair that co-occurred in the same document, originally from Yan et al (2013). Similarly, Yang et al (2020); Zhou et al (2020) use a bipartite graph of documents and words, connected by word occurrences or TF-IDF values. Wang et al (2021b) use word co-occurrence and semantic correlation graphs. Wang et al (2022b) focus on graph topic modeling with micro-blogs. Zhu et al (2023) propose a graph neural topic model to incorporate commonsense knowledge.

3.1.5 NTMs with generative adversarial networks

Some studies focus on employing Generative Adversarial Networks (GANs) to facilitate topic modeling. Wang et al (2019) follows the idea of GAN: they use a generator to generate “fake” documents from a random Dirichlet sample and then use a discriminator to distinguish the generated documents from real ones. Note that their model cannot infer doc-topic distributions because it directly maps documents to representations based on TF-IDF. To lift this limitation, Wang et al (2020) propose to use bidirectional adversarial training, which can meanwhile infer doc-topic distributions. Hu et al (2020) further present an extension that uses two cycle-consistency constraints to generate informative representations.

3.1.6 NTMs with pre-trained language models

Researchers frequently combine NTMs with pre-trained language models. Pre-trained language models based on Transformers (Vaswani et al 2017) have been prevalent in NLP fields, which are pre-trained on large-scale corpora to capture contextual linguistic features. Multiple studies leverage contextual features from these pre-trained models to provide richer information than conventional BoW. For instance, Bianchi et al (2021a) input the concatenation of BoW and the contextual document embeddings from Sentence-BERT (Reimers and Gurevych 2019), and then reconstruct BoW as previous work. Hoyle et al (2020) propose to distill knowledge from BERT (Devlin et al 2018) to NTMs. In detail, they produce pseudo BoW from the predictive word probability of BERT. Then their NTM reconstructs both the real and pseudo BoW. Bianchi et al (2021b); Mueller and Dredze (2021) employ multilingual BERT to infer cross-lingual doc-topic distributions for zero-shot learning but they cannot discover aligned cross-lingual topics.

3.1.7 NTMs with contrastive learning

As a self-supervised learning fashion, contrastive learning has been employed to facilitate NTMs. The idea of contrastive learning is to measure the similarity relations among sample pairs in a representation space (Van den Oord et al 2018). Nguyen and Luu (2021) propose the contrastive learning on doc-topic distributions where they build positive and negative pairs by sampling salient words of documents. Differently, Wu et al (2022) directly sample positive and negative pairs based on the topic semantics of documents to capture

relations among samples. Specifically, they quantize doc-topic distributions following Wu et al (2020b) and then sample documents with the same quantization indices as positive pairs and different indices as negative pairs. Their method can also capture the similarity relations among augmented samples. Zhou et al (2023) improve topic disentanglement with contrastive learning on word and topic embeddings. Han et al (2023) cluster documents, compute term weights, and make NTMs reconstruct salient words. They also use contrastive learning to refine doc-topic distributions where positive samples come from pre-trained language models.

3.1.8 NTMs with reinforcement learning

Reinforcement learning has been utilized to facilitate the learning process of NTMs. To be specific, Gui et al (2019) enhance NTMs with a reinforcement learning framework. They evaluate topic coherence performance during training and use this performance as reward signals to guide the learning of topic modeling. Costello and Reformat (2023) follow this idea and add more improvements like using sentence embeddings, adding a weighting term to the ELBO, and tracking topic diversity and coherence during training.

3.1.9 Other NTMs

Apart from the aforementioned methods, we introduce NTMs with other structures.

Before the invention of VAE-based NTMs, researchers have different attempts to model latent topics with neural networks. Some studies focus on NTMs in the autoregressive framework. Larochelle and Lauly (2012) propose an autoregressive NTM, called DocNADE. Inspired by Replicated Softmax (Hinton and Salakhutdinov 2009), DocNADE predicts the probability of a word in a document conditioned on its hidden state which is conditioned on previous words. Then it interprets topics with a hidden state and infers doc-topic distributions with the hidden states of the document. Gupta et al (2019b) extend DocNADE by modeling the bi-directional dependencies between words. Gupta et al (2019a) then use a LSTM to enable DocNADE to incorporate external knowledge.

Cao et al (2015) also propose an early NTM before VAE-based NTMs. Their approach predicts how an n-gram correlates with documents. It computes the representation of an n-gram by transforming the accumulation of the word embeddings from Word2Vec (Mikolov et al 2013) and projects documents into representations with a look-up matrix table. In this way, they model topic-word distributions as the n-gram representations and model doc-topic distributions as the projected document representations. For training, it uses the document of the n-gram as a positive and randomly samples documents that do *not* contain this n-gram as negatives.

Lin et al (2019) replace the softmax function with the sparsemax to enhance the sparsity in doc-topic distributions. Nan et al (2019) use Wasserstein AutoEncoder (WAE) to model topics, which minimizes the Wasserstein distance between generated documents and input documents. Rezaee and Ferraro (2020) propose a NTM without using the reparameterization trick. They generate discrete topic assignments from RNN inspired by Dieng et al (2017). Wu et al (2021) focus on discovering latent topics from long-tailed corpora. They propose a causal inference framework to analyze how the long-tailed bias influences topic modeling. Then they use a simple but effective casual intervention method to mitigate such influence.

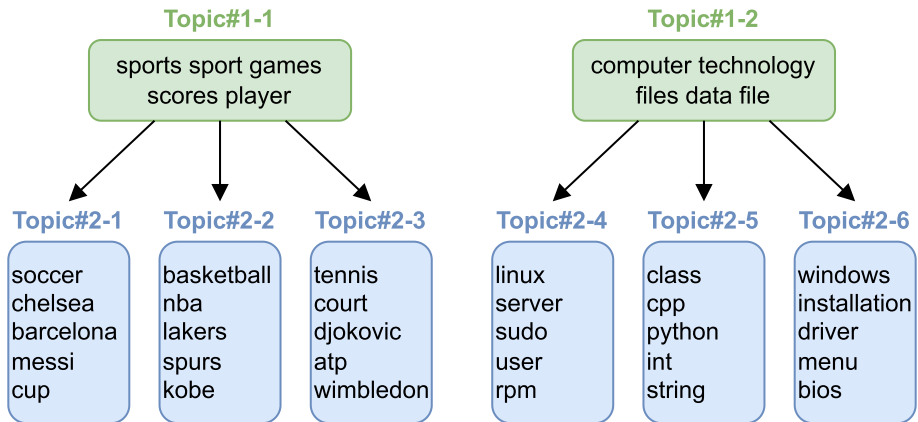


Fig. 4 Illustration of hierarchical topic modeling. Topics at each level cover different semantic granularity: child topics are more specific to parent topics. Topic#2-1 denotes the first topic of the second layer in the topic structure

3.1.10 Topic discovery by clustering

We discuss a special type of approach that discovers latent topics by clustering instead of modeling the generation process of documents. They typically leverage traditional word embeddings such as Word2Vec (Mikolov et al 2013) or contextual embeddings from pre-trained language models. **We must clarify that these methods differ from the aforementioned ordinary topic models.** This is because they can only produce topics but cannot infer doc-topic distributions as required. Accordingly, their one advantage is their enhanced computational efficiency. In detail, Thompson and Mimno (2020) straightforwardly cluster token-level word embeddings from pre-trained models like BERT and GPT-2 and produce topics from the words assigned to clusters. Similarly, Sia et al (2020); Angelov (2020); Zhang et al (2022c) cluster word or document embeddings and interpret hidden topics by sampling words from clusters via term weights like TF-IDF. Following this line of work, Grootendorst (2022) propose BERTopic by clustering document representations through HDBSCAN. Note that BERTopic can estimate the doc-topic distribution based on the term weights within a given document.

3.2 NTMs for various scenarios

Besides the basic scenario on normal documents, we in this section introduce NTMs tailored for various use case scenarios, such as hierarchical, cross-lingual, and dynamic topic modeling. We introduce the background of each scenario and present the related NTMs.

3.2.1 Hierarchical NTMs

Similar to conventional topic models (Griffiths et al 2003; Teh et al 2004; Blei et al 2010), NTMs can discover hierarchical topics to reveal topic structures from general to specific. Topics at each level in a hierarchy cover different semantic granularity: child topics tend

to be more specific to their parent topics. As shown in Fig. 4, a topic about “sports” can derive more specific child topics, like “soccer”, “basketball”, and “tennis”; a topic about “computer” also has specific child topics like “linux”, “programming”, and “windows”. In addition, hierarchical topic modeling can relieve the challenge of determining the number of topics to some extent (Blei et al 2010).

To discover hierarchical topics, some work follows the previous non-parametric setting. Isonuma et al (2020) propose a tree-structured neural topic model with two doubly-recurrent neural networks over the ancestors and siblings respectively (Alvarez-Melis and Jaakkola 2017). Note that the tree structure is unbounded, *i.e.*, it can be dynamically updated in a heuristic way during training. Pham and Le (2021) follow this spirit and jointly handle hierarchical topics and document visualization. Chen et al (2021b) leverage a stick-breaking process as prior for non-parametric modeling.

Later, the parametric fashion has attracted more attention, which requires setting the width and height of a topic hierarchy ahead. Chen et al (2021a) propose manifold regularization on topic hierarchy learning. Duan et al (2021) propose a Sawtooth Connection to model topic dependencies across hierarchical levels based on the model structure of ETM Dieng et al (2020). As aforementioned, Xu et al (2022) use different layers in the hyperbolic embedding space to interpret hierarchical topics. Li et al (2022) employ skip-connections for decoding to alleviate the posterior collapsing issue and propose a policy gradient method for training. Recently, Duan et al (2023) propose to generate different documents for different levels. They craft documents with more related words through word similarity matrices for higher levels, and then progressively generate these documents at each level. Chen et al (2023) utilize a Gaussian mixture prior and nonlinear structural equations to model topic dependencies between hierarchical levels. The main issue of the parametric setting is that topic hierarchies cannot grow dynamically since their width and height must be predetermined before training.

3.2.2 Short text NTMs

Researchers apply NTMs to discover topics from short texts. Short texts, prevalent on the Internet in various forms such as tweets, comments, and news headlines, serve as a common medium for individuals to express ideas, comments, and opinions. However, normal topic models often struggle to effectively handle short texts. The principal reason is that topic models depend on the word co-occurrence information to infer latent topics, but such information is extremely sparse in short texts due to their limited context. This challenge, referred to as *data sparsity* (Yan et al 2013; Wu and Li 2019), hinders topic models from discovering high-quality topics and thus has attracted considerable attention in the research community.

Several studies are proposed to overcome this data sparsity challenge. Lin et al (2020) use the Archimedean copulas to regularize the discreteness of topic distributions of short texts. Wu et al (2020b) propose to quantize doc-topic distributions of short texts to quantization vectors following the idea of Van den Oord and Vinyals (2017). By carefully initializing the quantization vectors, they can produce sharper doc-topic distributions that better fit short texts with limited context. They also propose a negative sampling decoder to avoid repetitive topics besides the negative log-likelihood. To address the data sparsity issue, Wang et al (2021b) use word co-occurrence and semantic correlation graphs to enrich the learning signals of short texts. Zhao et al (2021d) propose to incorporate entity vector representations into a NTM for short texts. They learn entity vector representations

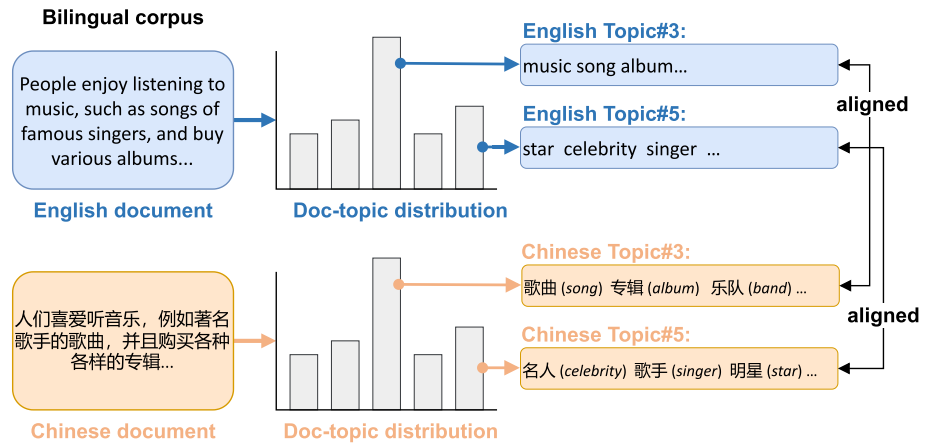


Fig. 5 Illustration of cross-lingual topic modeling (on English and Chinese documents). Corresponding cross-lingual topics are required to be aligned, like English and Chinese Topic#3, and English and Chinese Topic#5. *Words* in the brackets are the English translations

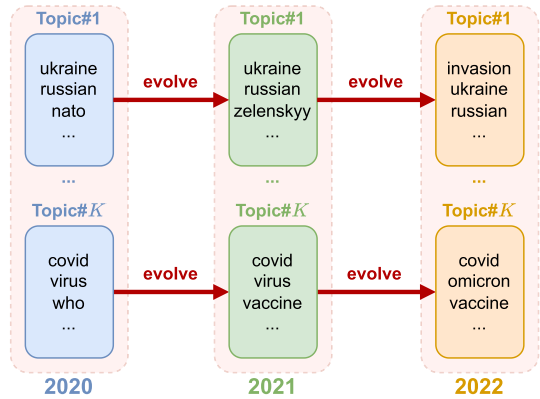
from manually edited knowledge graphs. Based on Wu et al (2020b), Wu et al (2022) further propose a contrastive learning method according to the topic semantics of short texts, which better captures the similarity relations among them. This refines the representations of short texts and thus their doc-topic distributions. They can also adapt to using data augmentation to further mitigate the data sparsity problem.

3.2.3 Cross-lingual NTMs

Cross-lingual NTMs are proposed following cross-lingual topic models (Mimno et al 2009). Cross-lingual topic models aim to discover aligned topics in different languages. As exemplified in Fig. 5, English and Chinese Topic#3 both refer to “music”, and English and Chinese Topic#5 refer to “celebrity”. In addition if two documents in different languages contain similar latent topics, their inferred doc-topic distributions should be similar. For instance, the doc-topic distributions of the parallel English and Chinese documents in Fig. 5 are similar. These aligned cross-lingual topics can reveal commonalities and differences across languages and cultures, which enables cross-lingual text analysis without supervision.

Wu et al (2020a) propose the first neural cross-lingual topic model. It transforms the topic-word distribution to the vocabulary space of another language. Thus the topic-word distributions of one language can incorporate the semantics of another language, which aligns cross-lingual topics. They show that their model outperforms traditional multilingual topic models (Shi et al 2016; Yuan et al 2018). Later, Wu et al (2023a) propose to align cross-lingual topics from the perspective of mutual information. This can properly align cross-lingual topics and prevent degenerate topic representations. To address the low-coverage dictionary issue, they also propose a cross-lingual vocabulary linking method that finds more linked words for topic alignment beyond the given dictionary. Bianchi et al (2021b); Mueller and Dredze (2021) directly learn cross-lingual doc-topic distributions with multilingual BERT. But we emphasize that they cannot discover aligned cross-lingual topics as required.

Fig. 6 Illustration of dynamic topic modeling. Topics associated with time slice t evolve from topics associated with slice $t - 1$. Here Topic#1 in 2022 evolves from Topic#1 in 2021. It is similar for other topics



3.2.4 Dynamic NTMs

Dynamic NTMs are explored following dynamic topic models (Blei and Lafferty 2006b; Wang et al 2012). Previous static topic models implicitly assume that documents are exchangeable. However, this assumption is inappropriate since documents are produced sequentially, such as scholarly journals, emails, and news articles. As such, dynamic topic models are proposed. While topics in previous methods are all static, dynamic topic models allow topics to shift over time to capture the topic evolution in sequential documents. To be specific, dynamic topic models assume that documents are divided by time slice, for example by year, and each time slice has K latent topics. The topics associated with slice t evolve from the topics associated with slice $t - 1$. As the example in Fig. 6, Topic#1 about Ukraine and Russia evolves from the year 2020 to 2022. Due to the emergence of the word “invasion”, we see Topic#1 captures the Ukraine-Russia war that exploded in 2022. Similarly, Topic#K about Covid-19 evolves from the year 2020 to 2022 with the explosion of the Omicron variant. These topic evolution reveals how topics emerge, grow, and vanish, which has been applied for trend analysis and opinion mining.

Dieng et al (2019) first propose a neural dynamic topic model, the Dynamic Embedding Topic Model. It uses word and topic embeddings to interpret latent topics following Dieng et al (2020). The topic embeddings at slice t depend on topic embeddings at slice $t - 1$. Besides, it uses a LSTM to learn temporal priors of doc-topic distributions. Rahimi et al (2023) discover topic evolution by clustering documents but cannot infer doc-topic distributions as required. Zhang and Lauw (2022) focus on the dynamic topics of temporal document networks and incorporate the linking information between documents. Rather than modeling topic evolution, Cvejovski et al (2023) model the activities of topics over time. Note that the activities of topics evolve over time but their topics are invariant. Thus, this method does *not* precisely adhere to the original definition of dynamic topic modeling.

3.2.5 Correlated NTMs

Following the idea of correlated topic modeling (Blei and Lafferty 2006a), correlated NTMs have been explored. Correlated topic models seek to consider the correlation between latent topics. For example, a document about genetics is more likely to be also about disease than x-ray astronomy (Blei and Lafferty 2006a). This leads to better

expressiveness than LDA. Liu et al (2019) follow the VAE-based NTM and use centralized transformation flow to capture topic correlations. To effectively infer the transformation flow, they present the transformation flow lower bound to regulate the KL divergence term.

3.2.6 Lifelong NTMs

Lifelong NTMs are proposed to solve the challenge of data sparsity, similar to short text NTMs but in a continual lifelong learning fashion. Gupta et al (2020) propose the first lifelong NTM. They retain prior knowledge, *i.e.*, topics, from document streams and guide topic modeling on sparse datasets with the accumulated knowledge. In detail, they use topic regularization to transfer topical knowledge from several domains and prevent catastrophic forgetting, and a selective replay strategy to identify relevant historical documents. Zhang et al (2022a) propose a lifelong NTM enhanced with a knowledge extractor and adversarial networks.

Although lifelong and dynamic topic modeling both work on sequential documents, we emphasize their difference: Dynamic topic modeling targets to discover topic evolution, *i.e.*, replacing outdated topics with emergent ones. Lifelong topic modeling aims to mitigate the data sparsity issue by accumulating prior topical knowledge, so it needs to restrain from forgetting past knowledge.

4 Applications of NTMs

In this section, we introduce the applications of NTMs, mainly including text analysis, text generation, and content recommendation.

4.1 Text analysis

The primary applications of NTMs concentrate on text analysis (Hoyle et al 2022; Laureate et al 2023).

Bai et al (2018) apply a NTM to analyze scientific articles. They enable a NTM to incorporate the citation graphs of scientific articles by predicting the connections between them. Thus their model can also recommend related articles to users. Zeng et al (2018) combine a NTM and a memory network and jointly train them for short text classification. Their method classifies short texts and discovers topics from them simultaneously. Chaudhary et al (2020) combine BERT with a NTM, which reduces the operation of self-attention. They claim that this can greatly speed up their fine-tuning process and thus reduce CO₂ emission. Song et al (2021) propose a classification-aware NTM which includes a NTM and a classifier. They focus on classifying the disinformation about COVID-19 to help deliver effective public health messages.

Zeng et al (2019) apply NTMs to understand the discourse in micro-blog conversations. Li et al (2020) use a dynamic NTM to understand the global impact of COVID-19 and non-pharmacological interventions in different countries and media sources. Their discovered dynamic topics help researchers understand the progression of the epidemic. Valero et al (2022) propose a short text NTM for podcast short-text metadata. Gui et al (2020) propose a multitask mutual learning framework for sentiment analysis and topic detection. They make the topic-word distributions similar to the word-level attention vectors through

mutual learning. Avasthi et al (2022) use NTMs to mine topics from large-scale scientific and biomedical text corpora.

4.2 Text generation

Several studies apply NTMs to text generation tasks. Specifically, Tang et al (2019) propose a text generation model that learns semantics and structural features through a VAE-based NTM. Yang et al (2021) leverage NTMs to alleviate the information sparsity issue in long story generation. They map a short text to a low dimensional doc-topic distribution, from which they sample interrelated words as a skeleton. With the short text and the skeleton as input, they use a Transformer to generate long stories. Nguyen et al (2021) use the doc-topic distributions of NTM to enrich and control the global semantics for text summarization. Zhang et al (2022b) propose a neural hierarchical topic model to discover hierarchical topics from documents. Then they generate keyphrases under the hierarchical topic guidance.

4.3 Content recommendation

Similar to early work (Wang and Blei 2011), NTMs can cooperate with recommendation systems. Esmaili et al (2019) combines a NTM with a recommender system for reviews through a structured auto-encoder. Xie et al (2021) use a graph NTM for citation recommendation.

5 Challenges of NTMs

Despite their popularity, NTMs encounter several challenges. In this section, we conclude these main challenges as possible future research directions.

5.1 Lacking reliable evaluation

Inheriting from conventional topic models, the critical challenge of NTMs primarily lies in the lack of reliable evaluation. Current evaluation methods have been developed for years, but they have the following issues.

5.1.1 Absence of standard evaluation metrics

The topic modeling field lacks standard evaluation metrics. Resorting to human judgment provides one effective way to evaluate topic models, such as topic rating and word intrusion tasks (Lau et al 2014). Unfortunately, its reliance on human raters renders it expensive and time-consuming, limiting its feasibility for wide-scale comparisons. Owing to this, researchers generally depend on automatic evaluation metrics, such as the topic coherence and diversity mentioned in Sect. 2.2. However, these automatic metrics encounter the following two problems:

- **Inconsistent usage and settings of automatic metrics.** The usage and settings of automatic metrics vary across papers and even within a paper. For example, variations

Table 1 Examples of trivial and repetitive topics.

| | Top words |
|----------|--|
| | Trivial topics |
| Topic#1: | Just even like one go really come good away everything |
| Topic#2: | Abstract accept many long adding bad displayed good additional great |
| Topic#3: | Fact even really pretty seem seems nothing thing usually often |
| | Repetitive Topics |
| Topic#4: | <u>Sports</u> scores games <u>soccer</u> league tennis ncaa players <u>football</u> club |
| Topic#5: | <u>Sports</u> tennis <u>soccer</u> <u>hockey</u> games <u>football</u> beach match <u>players</u> ncaa |
| Topic#6: | <u>Sports</u> match cup <u>hockey</u> olympic <u>football</u> <u>players</u> sport <u>league</u> <u>soccer</u> |

Each row represents the top words of a topic. Trivial topics include less informative concepts; repetitive topics contain repeating words. Repetitions are underlined

include the number of top words, the number of topics, reference corpora, and coherence or diversity metrics. Consequently, the results are often confined to specific studies, impeding the comparability of NTMs across different research papers. Such inconsistencies have led some benchmarking studies to argue that the conventional LDA can still outperform NTMs in certain aspects (Doan and Hoang 2021; Hoyle et al 2022).

- **Questionable agreement between automatic metrics and human judgment.** Some investigations have revealed the discrepancies between the coherence metrics and human evaluation: they find that automatic metrics declare winner models when the corresponding human evaluation does not. This raises concerns that coherence metrics, originally designed for older models, possibly are incompatible with the newer neural topic models (Doogan and Buntine 2021; Hoyle et al 2021). We believe similar concerns may extend to diversity metrics: they may also be inconsistent with human assessments. We detail the reasons and offer a heuristic solution in Sect. 6.

Owing to the above problems, researchers appeal to explore automatic metrics that better approximate the preferences of real-world topic model users (Hoyle et al 2021; Stambach et al 2023). Thus, proposing standard and practical evaluation metrics is a promising and urgent future research direction for topic modeling.

5.1.2 Lacking standardized dataset pre-processing settings

The topic modeling field lacks standardized dataset pre-processing settings for topic model comparisons. Researchers routinely pre-process datasets before running topic models, like removing less frequent words and stop words. Recent studies find that different dataset pre-processing settings greatly impact topic modeling outcomes, such as the minimum and maximum document frequency, maximum vocabulary size, and stop word sets (Card et al 2018; Wu et al 2023b). However, these pre-processing settings vary substantially across papers even if they use the same benchmark datasets like 20newsgroup. These variations raise questions about the generalization ability of their methods across different pre-processing settings. Thus their claimed performance improvements may be untenable. In consequence, establishing standardized dataset pre-processing settings emerges as an imperative prerequisite for ensuring reliable and consistent evaluations of topic models.

5.2 Low-quality topics

Regardless of the simplification and popularity of NTMs, the quality of their discovered topics has been questioned from two aspects:

- **Trivial Topics:** Discovered topics are trivial with uninformative words. These topics cannot reveal the actual latent semantics of documents. As exemplified in Table 1, the topics include “even”, “just”, and “really”. It is difficult to discern their underlying conceptual semantics.
- **Repetitive Topics:** Discovered topics are repetitive with the same words, also referred to as the topic collapsing problem. As shown in Table 1, the topics include the same words like “sports”, “games”, and “soccer”. It is hard to distinguish them. Apart from that, these repetitive topics imply some semantics are still hidden in documents.

More disastrously, some NTMs may exhibit triviality and repetitiveness in their discovered topics simultaneously (Wu et al 2020b, 2023b). These two kinds of low-quality topics impede the understanding, undermine the interpretability of topic models, and are less beneficial for downstream tasks and applications. In consequence, how to effectively and consistently overcome this challenge becomes a necessary and constructive research direction.

5.3 Sensitivity to hyperparameters

Another significant challenge of NTMs lies in their sensitivity to hyperparameters. Due to the complicated structures, NTMs typically possess more hyperparameters compared to conventional topic models. For example, hyperparameters such as dropout probability, batch size, and learning rate assume critical roles in several NTMs (Srivastava and Sutton 2017; Card et al 2018). Besides, certain NTMs cannot perform well under a large number of topics (Wu et al 2020b). As a result, researchers must meticulously fine-tune these hyperparameters when applying NTMs, especially to new datasets. Therefore, the sensitivity of NTMs to hyperparameters curtails the generalization ability of NTMs, underscoring the necessity to mitigate this sensitivity.

6 Topic semantic-aware diversity

In this section, we propose a new diversity metric that considers the semantics of topics when measuring topic diversity.

6.1 Problem of previous diversity metrics

Previous topic diversity metrics may contradict human judgment. These diversity metrics, such as TR (Burkhardt and Kramer 2019), TU (Nan et al 2019), and TD (Dieng et al 2020), all consider the uniqueness of one top word of topics. They believe that diversity is perfect only when each top word is unique. However, we argue that this measurement is over-strict since it ignores the fact that different topics may naturally share the same words due to word polysemy. As the examples shown in Table 2, “apple” refers to a kind of fruit in Topic#1 and a technology company in Topic#2, and “jobs” refers to Steve Jobs in Topic#2 or a paid position of employment in Topic#3. These topics imply different conceptual

Table 2 Comparison of topic diversity metrics under three cases

| Top words | TU | TD | TSD |
|--|-------|-------|-------|
| Case 1 | | | |
| Topic#1: <u>apple</u> peach grape orange banana | 0.867 | 0.733 | 1.000 |
| Topic#2: <u>apple</u> company steve <u>jobs</u> macintosh | | | |
| Topic#3: <u>jobs</u> unemployment economy worker salary | | | |
| Case 2 | | | |
| Topic#1: <u>apple</u> peach grape <u>orange</u> banana | 0.800 | 0.600 | 0.933 |
| Topic#2: <u>apple</u> <u>orange</u> steve <u>jobs</u> macintosh | | | |
| Topic#3: <u>jobs</u> unemployment economy worker salary | | | |
| Case 3 | | | |
| Topic#1: <u>apple</u> peach <u>grape</u> <u>orange</u> banana | 0.333 | 0.000 | 0.000 |
| Topic#2: <u>apple</u> peach <u>grape</u> <u>orange</u> banana | | | |
| Topic#3: <u>apple</u> peach <u>grape</u> <u>orange</u> banana | | | |

TU (Equation (5)) and TD (Equation (7)) refer to previous diversity metrics in Sect. 2.2, and TSD (Equation (16)) is our proposed new diversity metric. Each row represents the top words of a topic. Repetitive words are underlined

semantics although they include some same words. So we conceive their diversity score should be highest. But we see that their TU score is only 0.867 and TD is 0.733 in Table 2, which disagrees with our judgment.

6.2 Topic semantic-aware diversity

To address this issue, we propose **Topic Semantic-aware Diversity (TSD)**, a new metric that measures topic diversity along with word semantics.

6.2.1 Definition of topic semantic-aware diversity

In detail, we compute TSD based on the frequencies of word pairs. Given K topics and the top T words of each topic, we propose the new TSD as follows:

$$TSD = \frac{2}{KT(T-1)} \sum_{k=1}^K \sum_{(x_i, x_j) \in t(k)} \mathbb{I}(\#(x_i, x_j)). \tag{16}$$

Here $\#(x_i, x_j)$ means the number of an unordered word pair (x_i, x_j) in the top words of all K topics. $\mathbb{I}(\cdot)$ refers to an indicator function that equals to 1 if $\#(x_i, x_j) = 1$ and equals 0 otherwise. $t(k)$ denotes the top words of k -th topic. Rather than the uniqueness of one word, our TSD measures the uniqueness of word pairs in the top words of topics. This is because we know what a word exactly refers to when paired with another one. For example, “apple” refers to fruit if paired with “orange” or “banana” and to a company if with “technology” or “company”. Note that TSD degrades to TD when measuring the frequency of each word in Equation (16).

We exemplify the difference between our TSD with previous diversity metrics. Table 2 Case 1 shows the TSD score of these three topics is 1.0. This is because “apple” does *not*

Table 3 Correlations between topic diversity metrics and human ratings on different datasets

| Datasets | TU | TD | TSD |
|-------------|-------|-------|-------|
| NeurIPS | 0.970 | 0.970 | 0.984 |
| ACL | 0.969 | 0.970 | 0.998 |
| NYT | 0.913 | 0.913 | 0.963 |
| Wikitext103 | 0.973 | 0.973 | 0.999 |
| Average | 0.957 | 0.957 | 0.986 |

co-occur with “orange”, “grape”, or “banana”, and “jobs” does *not* co-occur with “unemployment”, “economy”, or “salary” in Topic#2. Thus TSD considers Topic#1-3 as different topics regardless of the same words. Naturally, TSD punishes diversity if the word pairs are repetitive. For instance, Table 2 Case 2 shows the TSD score of the three topics becomes lower since “apple” co-occurs with “orange” in both Topic#1 and #2. In the worst situation, Table 2 Case 3 has all the same topics. We see in this case both TD and TSD give 0 for topic diversity.

6.2.2 Evaluation results

We conduct experiments to sufficiently compare our proposed topic diversity metric and previous ones. In detail, we employ a conventional topic model, LDA (Blei et al 2003) and a neural topic model, NSTM (Zhao et al 2021c) to discover latent topics from real-world datasets. Then we ask human raters to evaluate the diversity among the top words of sampled topics. The adopted datasets are listed as follows: (i) NeurIPS⁶, including published papers at the NeurIPS conference from 1987 to 2017. (ii) ACL (Bird et al 2008), including research articles between 1973 and 2006 from the ACL Anthology.⁷ (iii) NYT,⁸ including news articles on the New York Times website from 2012 to 2022. (iv) Wikitext103⁹ (Merity et al 2016), including Wikipedia articles. Following Lau et al (2014); Röder et al (2015), we compute the Pearson correlation coefficients between the results of these topic diversity metrics and human ratings.

Table 3 shows the correlation results on different datasets and the average correlation. We notice that our TSD achieves relatively higher correlation scores with human ratings. This is because our TSD metric considers the word semantics as well while measuring topic diversity. These empirical results demonstrate that our TSD metric more closely aligns with human judgment concerning the topic diversity evaluation.

⁶ <https://www.kaggle.com/datasets/benhamner/nips-papers>.

⁷ <https://aclanthology.org/>.

⁸ https://huggingface.co/datasets/Matthewww/nyt_news.

⁹ <https://www.salesforce.com/products/einstein/ai-research/the-wikitext-dependency-language-modeling-dataset/>.

7 Topic model toolkits

Several topic model toolkits have been developed by the research community. Early popular toolkits include MALLET¹⁰ (McCallum 2002), gensim¹¹ (Rehurek and Sojka 2011), STTM¹² (Qiang et al 2020), ToModAPI¹³ (Lisena et al 2020), and tomotopy¹⁴. However these toolkits often neglect either the implementations of NTMs, dataset pre-processing, or evaluations, leaving a gap in meeting practical requirements. Recently, Terragni et al (2021) propose the OCTIS toolkit¹⁵ which includes several NTM methods, evaluations, and Bayesian parameter optimization for research. The latest toolkit is TopMost¹⁶ (Wu et al 2023c). Compared to OCTIS, TopMost covers a wider range of topic modeling scenarios and more newest released NTMs. It also decouples the model implementations and training, which eases the extension of new models. These toolkits provide a solid foundation for beginners to explore various topic models and empower users to apply diverse topic models in their applications.

8 Conclusion

Topic models have been prevalent for decades with diverse applications. Recently Neural Topic Models (NTMs) have attracted significant attention due to their flexibility and scalability. They stand out by offering advantages such as avoiding the requirement for model-specific derivations and efficiently handling large-scale datasets. With the emergence of NTMs, researchers have explored several promising applications for various tasks.

In this paper, we provide a comprehensive and up-to-date survey of NTMs. We introduce the preliminary of topic modeling, including the problem setting, notations, and evaluation methods. We review the existing NTM methods that employ different network structures and discuss their applicability to different use case scenarios. In addition, we delve into an examination of the popular applications built on NTMs. Finally, we identify and discuss the challenges that lie ahead for NTM research in detail. We hope this survey can serve as a valuable resource for researchers interested in NTMs and contribute to the advancement of NTM research.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

¹⁰ <https://mimno.github.io/Mallet/topics.html>.

¹¹ <https://radimrehurek.com/gensim/>.

¹² <https://github.com/qiang2100/STTM>.

¹³ <https://github.com/D2KLab/ToModAPI>.

¹⁴ <https://github.com/bab2min/tomotopy>.

¹⁵ <https://github.com/MIND-Lab/OCTIS/>.

¹⁶ <https://github.com/bobxwu/topmost>.

References

- Alvarez-Melis D, Jaakkola TS (2017) Tree-structured decoding with doubly-recurrent neural networks. In: International Conference on Learning Representations
- Angelov D (2020) Top2vec: Distributed representations of topics. arXiv preprint [arXiv:2008.09470](https://arxiv.org/abs/2008.09470)
- Avasthi S, Chauhan R, Acharjya DP (2022) Topic modeling techniques for text mining over a large-scale scientific and biomedical text corpus. *Int J Ambient Comput Intell* 13(1):1–18
- Bai H, Chen Z, Lyu MR, et al (2018) Neural relational topic models for scientific article analysis. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, pp 27–36
- Bianchi F, Terragni S, Hovy D (2021a) Pre-training is a hot topic: Contextualized document embeddings improve topic coherence. In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers), pp 759–766
- Bianchi F, Terragni S, Hovy D, et al (2021b) Cross-lingual contextualized topic models with zero-shot learning. In: Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume. Association for Computational Linguistics, Online, pp 1676–1683, <https://doi.org/10.18653/v1/2021.eacl-main.143>, <https://aclanthology.org/2021.eacl-main.143>
- Bird S, Dale R, Dorr BJ, et al (2008) The acl anthology reference corpus: a reference dataset for bibliographic research in computational linguistics. In: LREC
- Blei D, Lafferty J (2006) Correlated topic models. *Adv Neural Inf Process Syst* 18:147
- Blei DM (2012) Probabilistic topic models. *Commun ACM* 55(4):77–84
- Blei DM, Lafferty JD (2006b) Dynamic topic models. In: Proceedings of the 23rd international conference on Machine learning, pp 113–120
- Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet allocation. *J Mach Learn Res* 3:993–1022
- Blei DM, Griffiths TL, Jordan MI (2010) The nested Chinese restaurant process and Bayesian nonparametric inference of topic hierarchies. *J ACM* 57(2):1–30
- Blei DM, Kucukelbir A, McAuliffe JD (2017) Variational inference: a review for statisticians. *J Am Stat Assoc* 112(518):859–877
- Bouma G (2009) Normalized (pointwise) mutual information in collocation extraction. In: Proceedings of GSCL, pp 31–40
- Boyd-Graber JL, Hu Y, Mimno D et al (2017) Applications of topic models, vol 11. Springer, New York
- Buntine WL (2009) Estimating likelihoods for topic models. *ACML* 9:51–64
- Burkhardt S, Kramer S (2019) Decoupling sparsity and smoothness in the Dirichlet variational autoencoder topic model. *J Mach Learn Res* 20(131):1–27
- Cao Z, Li S, Liu Y, et al (2015) A novel neural topic model and its supervised extension. In: Proceedings of the AAAI Conference on Artificial Intelligence
- Card D, Tan C, Smith NA (2018) Neural Models for Documents with Metadata. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp 2031–2040
- Chang J, Gerrish S, Wang C, et al (2009) Reading tea leaves: How humans interpret topic models. In: Advances in neural information processing systems, pp 288–296
- Chaudhary Y, Gupta P, Saxena K et al (2020) Topicbert for energy efficient document classification. *Find Assoc Comput Ling* 2020:1682–1690
- Chen H, Mao P, Lu Y, et al (2023) Nonlinear structural equation model guided gaussian mixture hierarchical topic modeling. In: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp 10377–10390
- Chen Z, Ding C, Rao Y et al (2021) Hierarchical neural topic modeling with manifold regularization. *World Wide Web* 24:2139–2160
- Chen Z, Ding C, Zhang Z, et al (2021b) Tree-structured topic modeling with nonparametric neural variational inference. In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pp 2343–2353
- Churchill R, Singh L (2022) The evolution of topic modeling. *ACM Comput Surv* 54(10s):1–35
- Costello J, Reformat MZ (2023) Reinforcement learning for topic models. arXiv preprint [arXiv:2305.04843](https://arxiv.org/abs/2305.04843)
- Cvejovski K, Sánchez RJ, Ojeda C (2023) Neural dynamic focused topic model. arXiv preprint [arXiv:2301.10988](https://arxiv.org/abs/2301.10988)
- Devlin J, Chang MW, Lee K, et al (2018) Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint [arXiv:1810.04805](https://arxiv.org/abs/1810.04805)

- Dieng AB, Wang C, Gao J, et al (2017) TopicRNN: A recurrent neural network with long-range semantic dependency. In: International Conference on Learning Representations, <https://openreview.net/forum?id=rJbbOLcex>
- Dieng AB, Ruiz FJ, Blei DM (2019) The dynamic embedded topic model. arXiv preprint [arXiv:1907.05545](https://arxiv.org/abs/1907.05545)
- Dieng AB, Ruiz FJ, Blei DM (2020) Topic modeling in embedding spaces. *Trans Assoc Comput Lingu* 8:439–453
- Doan TN, Hoang TA (2021) Benchmarking neural topic models: An empirical study. In: Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021. Association for Computational Linguistics, Online, pp 4363–4368, 10.18653/v1/2021.findings-acl.382, <https://aclanthology.org/2021.findings-acl.382>
- Doogan C, Buntine W (2021) Topic model or topic twaddle? Re-evaluating semantic interpretability measures. In: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp 3824–3848
- Duan Z, Wang D, Chen B, et al (2021) Sawtooth factorial topic embeddings guided gamma belief network. In: International Conference on Machine Learning, PMLR, pp 2903–2913
- Duan Z, Xu Y, Sun J, et al (2022) Bayesian deep embedding topic meta-learner. In: International Conference on Machine Learning, PMLR, pp 5659–5670
- Duan Z, Liu X, Su Y, et al (2023) Bayesian progressive deep topic model with knowledge informed textual data coarsening process. In: International Conference on Machine Learning, PMLR, pp 8731–8746
- Duong C, Liu Q, Mao R, et al (2022) Saving earth one tweet at a time through the lens of artificial intelligence. In: 2022 International Joint Conference on Neural Networks (IJCNN), pp 1–9, 10.1109/IJCNN55064.2022.9892271
- Esmacili B, Huang H, Wallace B, et al (2019) Structured neural topic models for reviews. In: Chaudhuri K, Sugiyama M (eds) Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research, vol 89. PMLR, pp 3429–3439, <https://proceedings.mlr.press/v89/esmacili19b.html>
- Griffiths T, Jordan M, Tenenbaum J, et al (2003) Hierarchical topic models and the nested Chinese restaurant process. *Adv Neural Inf Process Syst* 16
- Grootendorst M (2022) Bertopic: Neural topic modeling with a class-based tf-idf procedure. arXiv preprint [arXiv:2203.05794](https://arxiv.org/abs/2203.05794)
- Gui L, Leng J, Pergola G, et al (2019) Neural topic model with reinforcement learning. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pp 3478–3483
- Gui L, Leng J, Zhou J et al (2020) Multi task mutual learning for joint sentiment classification and topic detection. *IEEE Trans Knowl Data Eng* 34(4):1915–1927
- Gupta P, Chaudhary Y, Buettner F, et al (2019a) Texttavec: Deep contextualized neural autoregressive topic models of language with distributed compositional prior. In: International Conference on Learning Representations
- Gupta P, Chaudhary Y, Buettner F, et al (2019b) Document informed neural autoregressive topic models with distributional prior. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp 6505–6512
- Gupta P, Chaudhary Y, Runkler T, et al (2020) Neural topic modeling with continual lifelong learning. In: International Conference on Machine Learning, PMLR, pp 3907–3917
- Han S, Shin M, Park S, et al (2023) Unified neural topic model via contrastive learning and term weighting. In: Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, Dubrovnik, Croatia, pp 1802–1817, <https://aclanthology.org/2023.eacl-main.132>
- Hennig P, Stern P, Herbrich R, et al (2012) Kernel topic models. In: *Artificial Intelligence and Statistics*, pp 511–519
- Hinton GE, Salakhutdinov RR (2009) Replicated softmax: an undirected topic model. *Advances in Neural Information Processing Systems* vol 22
- Hoyle A, Goel P, Hian-Cheong A, et al (2021) Is automated topic model evaluation broken? the incoherence of coherence. In: Beygelzimer A, Dauphin Y, Liang P, et al (eds) *Advances in Neural Information Processing Systems*, <https://openreview.net/forum?id=tjdHCnPqoo>
- Hoyle AM, Goel P, Resnik P (2020) Improving neural topic models using knowledge distillation. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)
- Hoyle AM, Sarkar R, Goel P et al (2022) Are neural topic models broken? *Find Assoc Comput Ling* 2022:5321–5344

- Hu X, Wang R, Zhou D, et al (2020) Neural topic modeling with cycle-consistent adversarial training. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp 9018–9030
- Isonuma M, Mori J, Bollegala D, et al (2020) Tree-structured neural topic model. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp 800–806
- Joo W, Lee W, Park S et al (2020) Dirichlet variational autoencoder. *Pattern Recogn* 107:107514
- Kim H, Choo J, Kim J, et al (2015) Simultaneous discovery of common and discriminative topics via joint nonnegative matrix factorization. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp 567–576
- Kingma DP, Welling M (2014) Auto-encoding variational bayes. In: The International Conference on Learning Representations (ICLR)
- Korshunova I, Xiong H, Fedoryszak M, et al (2019) Discriminative topic modeling with logistic lda. In: Wallach H, Larochelle H, Beygelzimer A, et al (eds) *Advances in Neural Information Processing Systems*, vol 32. Curran Associates, Inc., https://proceedings.neurips.cc/paper_files/paper/2019/file/54ebdfbbf6c31c39aaba9a1ee83860a-Paper.pdf
- Krasnashchok K, Jouili S (2018) Improving Topic Quality by Promoting Named Entities in Topic Modeling. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, pp 247–253
- Larochelle H, Lauly S (2012) A neural autoregressive topic model. *Advances in Neural Information Processing Systems* 25
- Lau JH, Newman D, Baldwin T (2014) 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, pp 530–539
- Laureate CDP, Buntine W, Linger H (2023) A systematic review of the use of topic models for short text social media analysis. *Artificial Intelligence Review* pp 1–33
- Lee D, Seung HS (2000) Algorithms for non-negative matrix factorization. In: *Advances in Neural Information Processing Systems*, vol 13
- Li Y, Nair P, Wen Z, et al (2020) Global surveillance of covid-19 by mining news media using a multi-source dynamic embedded topic model. In: Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, pp 1–14
- Li Y, Wang C, Duan Z et al (2022) Alleviating “posterior collapse” in deep topic models via policy gradient. *Adv Neural Inf Process Syst* 35:22562–22575
- Lin L, Jiang H, Rao Y (2020) Copula guided neural topic modelling for short texts. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 1773–1776
- Lin T, Hu Z, Guo X (2019) Sparsemax and relaxed wasserstein for topic sparsity. In: Proceedings of the twelfth ACM international conference on web search and data mining, pp 141–149
- Lisena P, Harrando I, Kandakji O, et al (2020) Tomodapi: a topic modeling api to train, use and compare topic models. In: Proceedings of second workshop for NLP open source software (NLP-OSS), pp 132–140
- Liu L, Huang H, Gao Y, et al (2019) Neural variational correlated topic modeling. In: The World Wide Web Conference, pp 1142–1152
- Liu Z, Zhang Y, Chang EY et al (2011) Plda+ parallel latent Dirichlet allocation with data placement and pipeline processing. *ACM Trans Intell Syst Technol* 2(3):1–18
- van der Maaten L, Hinton G (2008) Visualizing data using t-SNE. *J Mach Learn Res* 9:2579–2605
- Manning CD, Raghavan P, Schütze H (2008) *Introduction to information retrieval*. Cambridge University Press, New York
- Mcauliffe J, Blei D (2007) Supervised topic models. *Adv Neural Inf Process Syst* 20:121–128
- McCallum AK (2002) Mallet: A machine learning for languagetoolkit. <http://mallet.cs.umass.edu>
- Merity S, Xiong C, Bradbury J, et al (2016) Pointer sentinel mixture models. arXiv preprint [arXiv:1609.07843](https://arxiv.org/abs/1609.07843)
- Miao Y, Yu L, Blunsom P (2016) Neural variational inference for text processing. In: International Conference on Machine Learning, pp 1727–1736
- Miao Y, Grefenstette E, Blunsom P (2017) Discovering discrete latent topics with neural variational inference. In: Proceedings of the 34th International Conference on Machine Learning-Volume 70. JMLR. org, pp 2410–2419
- Mikolov T, Chen K, Corrado G, et al (2013) Efficient estimation of word representations in vector space. In: Bengio Y, LeCun Y (eds) 1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings, [arXiv:abs/1301.3781](https://arxiv.org/abs/1301.3781)

- Mimno D, Wallach H, Naradowsky J, et al (2009) Polylingual topic models. In: Proceedings of the 2009 conference on empirical methods in natural language processing. Association for Computational Linguistics, Singapore, pp 880–889, <https://aclanthology.org/D09-1092>
- Mimno D, Wallach HM, Talley E, et al (2011) Optimizing semantic coherence in topic models. In: Proceedings of the conference on empirical methods in natural language processing. Association for Computational Linguistics, pp 262–272
- Mueller A, Dredze M (2021) Fine-tuning encoders for improved monolingual and zero-shot polylingual neural topic modeling. In: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp 3054–3068
- Nan F, Ding R, Nallapati R, et al (2019) Topic modeling with Wasserstein autoencoders. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, Florence, Italy, pp 6345–6381
- Newman D, Asuncion A, Smyth P, et al (2009) Distributed algorithms for topic models. *Journal of Machine Learning Research* 10(8)
- Newman D, Lau JH, Grieser K, et al (2010) Automatic evaluation of topic coherence. In: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, pp 100–108
- Nguyen T, Luu AT (2021) Contrastive learning for neural topic model. *Advances in Neural Information Processing Systems* 34
- Nguyen T, Luu AT, Lu T, et al (2021) Enriching and controlling global semantics for text summarization. In: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, pp 9443–9456
- Van den Oord A, Vinyals O (2017) Neural discrete representation learning. In: *Advances in Neural Information Processing Systems*, pp 6306–6315
- Van den Oord A, Li Y, Vinyals O (2018) Representation learning with contrastive predictive coding. *arXiv e-prints* pp arXiv-1807
- Pennington J, Socher R, Manning CD (2014) Glove: Global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 1532–1543
- Pham D, Le TM (2021) Neural topic models for hierarchical topic detection and visualization. In: *Machine Learning and Knowledge Discovery in Databases. Research Track: European Conference, ECML PKDD 2021, Bilbao, Spain, September 13–17, 2021, Proceedings, Part III 2021*, Springer, pp 35–51
- Qiang J, Qian Z, Li Y et al (2020) Short text topic modeling techniques, applications, and performance: a survey. *IEEE Trans Knowl Data Eng* 34(3):1427–1445
- Rahimi H, Naacke H, Constantin C, et al (2023) Antm: An aligned neural topic model for exploring evolving topics. *arXiv preprint arXiv:2302.01501*
- Rehurek R, Sojka P (2011) Gensim—statistical semantics in python. Retrieved from [gensim.org https://api.semanticscholar.org/CorpusID:64026679](https://api.semanticscholar.org/CorpusID:64026679)
- Reimers N, Gurevych I (2019) Sentence-bert: Sentence embeddings using siamese bert-networks. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pp 3982–3992
- Rezaee M, Ferraro F (2020) A discrete variational recurrent topic model without the reparametrization trick. *Adv Neural Inf Process Syst* 33:13831–13843
- Rezende DJ, Mohamed S, Wierstra D (2014) Stochastic backpropagation and approximate inference in deep generative models. In: Proceedings of the 31th International Conference on Machine Learning
- Röder M, Both A, Hinneburg A (2015) Exploring the space of topic coherence measures. In: Proceedings of the eighth ACM international conference on Web search and data mining, ACM, pp 399–408
- Shi B, Lam W, Bing L, et al (2016) Detecting common discussion topics across culture from news reader comments. In: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp 676–685
- Shi T, Kang K, Choo J, et al (2018) Short-text topic modeling via non-negative matrix factorization enriched with local word-context correlations. In: Proceedings of the 2018 World Wide Web Conference, International World Wide Web Conferences Steering Committee, pp 1105–1114
- Sia S, Dalmia A, Mielke SJ (2020) Tired of topic models? clusters of pretrained word embeddings make for fast and good topics too! In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics, Online, pp 1728–1736, <https://doi.org/10.18653/v1/2020.emnlp-main.135>, <https://aclanthology.org/2020.emnlp-main.135>
- Sievert C, Shirley K (2014) Ldavis: A method for visualizing and interpreting topics. In: Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces, pp 63–70

- Song X, Petrak J, Jiang Y et al (2021) Classification aware neural topic model for covid-19 disinformation categorisation. *PLoS ONE* 16(2):e0247086
- Srivastava A, Sutton C (2017) Autoencoding variational inference for topic models. In: 5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24–26, 2017, Conference Track Proceedings. OpenReview.net. <https://openreview.net/forum?id=BybtVK9lg>
- Stammach D, Zouhar V, Hoyle A, et al (2023) Re-visiting automated topic model evaluation with large language models. arXiv preprint [arXiv:2305.12152](https://arxiv.org/abs/2305.12152)
- Steyvers M, Griffiths T (2007) Probabilistic topic models. *Handb Latent Seman Anal* 427(7):424–440
- Tang H, Li M, Jin B (2019) A topic augmented text generation model: Joint learning of semantics and structural features. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pp 5090–5099
- Teh Y, Jordan M, Beal M, et al (2004) Sharing clusters among related groups: Hierarchical dirichlet processes. *Advances in neural information processing systems* 17
- Terragni S, Fersini E, Galuzzi BG, et al (2021) OCTIS: Comparing and optimizing topic models is simple! In: Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations. Association for Computational Linguistics, pp 263–270. <https://www.aclweb.org/anthology/2021.eacl-demos.31>
- Thompson L, Mimno D (2020) Topic modeling with contextualized word representation clusters. arXiv preprint [arXiv:2010.12626](https://arxiv.org/abs/2010.12626)
- Tian R, Mao Y, Zhang R (2020) Learning vae-lda models with rounded reparameterization trick. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp 1315–1325
- Valero FB, Baranes M, Epure EV (2022) Topic modeling on podcast short-text metadata. In: 44th European Conference on Information Retrieval (ECIR)
- Vaswani A, Shazeer N, Parmar N, et al (2017) Attention is all you need. *Advances in neural information processing systems* 30
- Wallach HM, Murray I, Salakhutdinov R, et al (2009) Evaluation methods for topic models. In: Proceedings of the 26th Annual International Conference on Machine Learning, pp 1105–1112
- Wang C, Blei DM (2011) Collaborative topic modeling for recommending scientific articles. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp 448–456
- Wang C, Blei D, Heckerman D (2008) Continuous time dynamic topic models. In: Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence, pp 579–586
- Wang C, Blei D, Heckerman D (2012) Continuous time dynamic topic models. arXiv preprint [arXiv:1206.3298](https://arxiv.org/abs/1206.3298)
- Wang D, Guo D, Zhao H, et al (2022a) Representing mixtures of word embeddings with mixtures of topic embeddings. In: International Conference on Learning Representations. <https://openreview.net/forum?id=IYMuTbGzjFU>
- Wang H, He R, Liu H, et al (2022b) Topic model on microblog with dual-streams graph convolution networks. In: 2022 International Joint Conference on Neural Networks (IJCNN), IEEE, pp 1–8
- Wang R, Zhou D, He Y (2019) Atm: Adversarial-neural topic model. *Inf Process Manag* 56(6):102098
- Wang R, Hu X, Zhou D, et al (2020) Neural topic modeling with bidirectional adversarial training. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, Online, pp 340–350. <https://doi.org/10.18653/v1/2020.acl-main.32>, <https://aclanthology.org/2020.acl-main.32>
- Wang X, Yang Y (2020) Neural topic model with attention for supervised learning. In: Chiappa S, Calandra R (eds) Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research, vol 108. PMLR, pp 1147–1156. <https://proceedings.mlr.press/v108/wang20c.html>
- Wang Y, Bai H, Stanton M, et al (2009) Plda: Parallel latent dirichlet allocation for large-scale applications. In: Algorithmic Aspects in Information and Management: 5th International Conference, AAIM 2009, San Francisco, CA, USA, June 15–17, 2009. Proceedings 5, Springer, pp 301–314
- Wang Y, Li X, Ouyang J (2021a) Layer-assisted neural topic modeling over document networks. In: IJCAI, pp 3148–3154
- Wang Y, Li X, Zhou X et al (2021) Extracting topics with simultaneous word co-occurrence and semantic correlation graphs: neural topic modeling for short texts. *Find Assoc Comput Ling* 2021:18–27
- Wu X, Li C (2019) Short Text Topic Modeling with Flexible Word Patterns. In: International Joint Conference on Neural Networks

- Wu X, Li C, Zhu Y, et al (2020a) Learning Multilingual Topics with Neural Variational Inference. In: International Conference on Natural Language Processing and Chinese Computing
- Wu X, Li C, Zhu Y, et al (2020b) Short text topic modeling with topic distribution quantization and negative sampling decoder. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), Online, pp 1772–1782
- Wu X, Li C, Miao Y (2021) Discovering topics in long-tailed corpora with causal intervention. In: Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021. Association for Computational Linguistics, Online, pp 175–185, 10.18653/v1/2021.findings-acl.15, <https://aclanthology.org/2021.findings-acl.15>
- Wu X, Luu AT, Dong X (2022) Mitigating data sparsity for short text topic modeling by topic-semantic contrastive learning. In: Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, Abu Dhabi, United Arab Emirates, pp 2748–2760, <https://aclanthology.org/2022.emnlp-main.176>
- Wu X, Dong X, Nguyen T, et al (2023a) Infotm: A mutual information maximization perspective of cross-lingual topic modeling. arXiv preprint [arXiv:2304.03544](https://arxiv.org/abs/2304.03544)
- Wu X, Dong X, Nguyen T, et al (2023b) Effective neural topic modeling with embedding clustering regularization. In: International Conference on Machine Learning, PMLR
- Wu X, Pan F, Luu AT (2023c) Towards the topmost: A topic modeling system toolkit. arXiv preprint [arXiv:2309.06908](https://arxiv.org/abs/2309.06908)
- Xie Q, Zhu Y, Huang J et al (2021) Graph neural collaborative topic model for citation recommendation. *ACM Trans Inf Syst* 40(3):1–30
- Xu Y, Wang D, Chen B, et al (2022) Hyperminer: Topic taxonomy mining with hyperbolic embedding. In: Koyejo S, Mohamed S, Agarwal A, et al (eds) *Advances in Neural Information Processing Systems*, vol 35. Curran Associates, Inc., pp 31557–31570
- Yan X, Guo J, Lan Y, et al (2013) A biterm topic model for short texts. In: Proceedings of the 22nd International Conference on World Wide Web, ACM, pp 1445–1456
- Yang L, Wu F, Gu J et al (2020) Graph attention topic modeling network. In: Proceedings of The Web Conference 2020, pp 144–154
- Yang Y, Pan B, Cai D, et al (2021) Topnet: Learning from neural topic model to generate long stories. In: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, pp 1997–2005
- Yin J, Wang J (2014) A dirichlet multinomial mixture model-based approach for short text clustering. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, pp 233–242
- Yuan M, Van Durme B, Ying JL (2018) Multilingual anchoring: Interactive topic modeling and alignment across languages. *Advances in neural information processing systems* 31
- Zeng J, Li J, Song Y, et al (2018) Topic memory networks for short text classification. In: Proceedings of the Conference on Empirical Methods in Natural Language Process
- Zeng J, Li J, He Y et al (2019) What you say and how you say it: joint modeling of topics and discourse in microblog conversations. *Trans Assoc Comput Ling* 7:267–281
- Zhang DC, Lauw H (2022) Dynamic topic models for temporal document networks. In: International Conference on Machine Learning, PMLR, pp 26281–26292
- Zhang H, Chen B, Guo D, et al (2018) WHAI: Weibull hybrid autoencoding inference for deep topic modeling. In: International Conference on Learning Representations, <https://openreview.net/forum?id=S1cZsf-RW>
- Zhang X, Rao Y, Li Q (2022) Lifelong topic modeling with knowledge-enhanced adversarial network. *World Wide Web* 25(1):219–238
- Zhang Y, Jiang T, Yang T, et al (2022b) Htkg: Deep keyphrase generation with neural hierarchical topic guidance. In: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 1044–1054
- Zhang Z, Fang M, Chen L, et al (2022c) Is neural topic modelling better than clustering? an empirical study on clustering with contextual embeddings for topics. In: NAACL 2022-2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference, p 3886
- Zhao H, Phung D, Huynh V, et al (2021a) Topic modelling meets deep neural networks: A survey. In: Zhou ZH (ed) Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21. International Joint Conferences on Artificial Intelligence Organization, pp 4713–4720, <https://doi.org/10.24963/ijcai.2021/638>, survey Track
- Zhao H, Phung D, Huynh V, et al (2021b) Neural topic model via optimal transport. In: International Conference on Learning Representations, <https://openreview.net/forum?id=Oos98K9Lv-k>

- Zhao H, Phung D, Huynh V, et al (2021c) Neural topic model via optimal transport. In: 9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3–7, 2021. OpenReview.net, <https://openreview.net/forum?id=Oos98K9Lv-k>
- Zhao X, Wang D, Zhao Z et al (2021) A neural topic model with word vectors and entity vectors for short texts. *Inf Process Manag* 58(2):102455
- Zhou D, Hu X, Wang R (2020) Neural topic modeling by incorporating document relationship graph. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp 3790–3796
- Zhou X, Bu J, Zhou S et al (2023) Improving topic disentanglement via contrastive learning. *Inf Process Manag* 60(2):103164
- Zhu B, Cai Y, Ren H (2023) Graph neural topic model with commonsense knowledge. *Inf Process Manag* 60(2):103215
- Zhu Q, Feng Z, Li X (2018) Graphbtm: Graph enhanced autoencoded variational inference for biterm topic model. In: Conference on Empirical Methods in Natural Language Processing (EMNLP 2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.