

Dynamic Topic Modeling by Clustering Embeddings from Pretrained Language Models: A Research Proposal

Anton Eklund
Umeå University
Adlede AB
Umeå, Sweden

anton.eklund@cs.umu.se

Mona Forsman
Adlede AB
Umeå, Sweden

mona.forsman@adlede.com

Frank Drewes
Umeå University
Umeå, Sweden

drewes@cs.umu.se

Abstract

A new trend in topic modeling research is to do Neural Topic Modeling by Clustering document Embeddings (NTM-CE) created with a pretrained language model. Studies have evaluated static NTM-CE models and found them performing comparably to, or even better than other topic models. An important extension of static topic modeling is making the models dynamic, allowing the study of topic evolution over time, as well as detecting emerging and disappearing topics. In this research proposal, we present two research questions to understand dynamic topic modeling with NTM-CE theoretically and practically. To answer these, we propose four phases with the aim of establishing evaluation methods for dynamic topic modeling, finding NTM-CE-specific properties, and creating a framework for dynamic NTM-CE. For evaluation, we propose to use both quantitative measurements of coherence and human evaluation supported by our recently developed tool.

1 Introduction

The ever-accelerating pace at which online documents, and specifically text documents, are published creates a need for methods able to analyze text documents in large quantities, something topic models were created to do. In this paper, the term *document* refers to a sequence of words in natural language, such as a tweet or a news article. The topic model analyzes a collection of documents to discover the major topics appearing in it. Then, each document is related to the discovered topics. Successful topic modeling applications cover a wide range of fields, such as studying historical documents (Newman and Block, 2006), discovering gender bias in datasets (Devinney et al., 2020), and catching new trends on social networks (Cataldi et al., 2010).

Topic models that do not consider the temporal dimension of a document collection are called

static topic models. However, an important aspect missed by such models is the evolution of topics or different temporal contexts a document can be situated in. For example, a topic centered around diabetes will have a different discussion before and after the discovery of insulin, and the topics surrounding *Ukraine* have shifted rather dramatically during 2022. Models that consider the temporal dimension are called *dynamic* topic models (DTMs). This paper proposes an extensive study on how to efficiently create DTMs based on neural topic models.

Neural Topic Models (NTMs) are topic models that are created with the help of neural networks (Zhao et al., 2021). They became competitive with the advances in language modeling in the previous decade. An NTM with an incorporated pretrained language model (PLM) is called NTM-PLM. A pretrained language model, such as BERT (Devlin et al., 2019), is capable of embedding words or documents into a vector representation that reflects aspects of the meaning of the text and thus its relation to other texts. We define a document embedding to be a mapping of the document collection to a vector space in such a way that the vector representing a given document captures some sort of information about the document. The information should largely be obtained using the PLM, but can also blend with other information if there is extra input such as timestamps. Since each two embedding vectors are at a certain well-defined distance from each other in the vector space, the conceptually most straightforward approach to do topic modeling is to apply a distance-based clustering algorithm to the embeddings. In this paper, we call this procedure *Neural Topic Modeling by Clustering Embeddings (from pretrained language models)*, NTM-CE.

The distinction between our definition of NTM-CE and other models within the NTM-PLM sphere is how the topics are formed. For a model to be an

NTM-CE, the topics must be formed by applying a distance-based clustering algorithm to the embeddings that were created by the PLM. We define the core pipeline of NTM-CE as *vectorization* \rightarrow *transformation* \rightarrow *clustering*. The most common NTM-CE methods that we are aware of use a dimension reduction technique as a vector transformation before clustering. However, we also consider other transformations that could be applied to enhance the vector space to discover meaningful topics. Belonging to our definition of NTM-CE models are CETopic (Zhang et al., 2022) and BERTopic (Groendorst, 2022), but not models such as *Embedding Topic Model* (ETM) (Dieng et al., 2020) or *ZeroShotTM* (Bianchi et al., 2021b) since the latter do not directly cluster the embeddings. NTM-CE models have shown promising performance (Sia et al., 2020; Thompson and Mimno, 2020; Zhang et al., 2022) when compared to classic generative methods such as *Latent Dirichlet Allocation* (LDA, Blei et al. (2003)).

Moreover, NTM-CE are gaining traction due to their conceptual simplicity and modularity. Early studies compare the topic coherence of NTM-CE with other established models to legitimize its use. However, an understudied part of these new topic models is their ability to do dynamic topic modeling. Having a conceptually simple DTM which improves with advancements in language modeling is attractive for many research communities and industries whose data tends to end up consisting of large unstructured sets of documents collected over time. Dynamic topic modeling may, for instance, enable a company to discover and react to trends before they become mainstream.

An important aspect of research on DTMs is how to evaluate such models. For the evaluation of static models, there has been substantial research on how different quantitative measurements relate to topic quality and human judgment (Chang et al., 2009; Lau et al., 2014). In contrast, little research has been devoted to the fair comparison of DTMs. Therefore, we suggest to develop a framework to make such comparisons. In this proposal, we present some potential ways in which such a comparative evaluation of DTMs could be done.

While NTM-CE is a promising technique, there is little research on adding a temporal dimension to NTM-CE and how to fairly compare DTMs. Therefore, we propose the following research questions in an attempt to thoroughly investigate the prospect

of using NTM-CE for dynamic topic modeling.

RQ1: What requirements exist for a dynamic topic modeling system and how can NTM-CE properties respond to those requirements?

RQ2: How viable is dynamic topic modeling with NTM-CE in practice?

With RQ1 we aim to lay a theoretical foundation that will lead to knowledge that remains valid beyond the current state of the art in language modeling. The goal is to create a general framework and to thoroughly address the strengths and weaknesses of NTM-CE from a dynamic perspective. RQ2 puts theoretical knowledge into practice and will reveal insights and limitations of dynamic NTM-CE. Here, we strive to create functional models that can solve problems of practical relevance for academia and industry.

2 Literature Review

Topic modeling is a field within text mining whose objective is to find topics that best describe a collection of documents and then assign the documents to these topics. Models from the stochastic school use Latent Dirichlet Allocation (LDA) (Blei et al., 2003) and its variants to fulfill this objective using probabilistic distributions to discover topics. The field of dynamic topic modeling (DTM¹) is the study of topics evolving over time. Discrete DTM such as Discrete LDA (d-LDA) (Blei and Lafferty, 2006) divides a topic into batches of discrete time steps where the next time step in a topic evolves from the previous. Continuous DTM (c-DTM) as introduced by Wang et al. (2008) borrowed the concept of Brownian motion from physics, which makes it possible to view the evolution of topics in continuous time. Another model that works continuously is *Topics Over Time* (TOT) by Wang and McCallum (2006) which associates each topic with a beta distribution representing the temporal dimension. In our project, we wish to study how to create a continuous-time topic model with NTM-CE as its basis. In addition, we want to use the dynamic models mentioned above for benchmarking future models.

With the introduction of Word2Vec (Mikolov et al., 2013), topic modeling saw a new era of models that make use of embeddings. The surveys by Zhao et al. (2021) and Churchill and Singh

¹We use the abbreviation DTM for both the modeling techniques and the resulting models.

(2021) describe how topic modeling meets neural networks in the modern era. They also describe many models that incorporate PLMs. Dieng et al. (2020) proposed a model called *Embedded Topic Model* (ETM) that merges LDA and Word2Vec into the same vector space. This makes it possible to extract more meaningful topics. The dynamic version of this model (d-ETM, Dieng et al. (2019)) similarly merges d-LDA with Word2Vec. Bianchi et al. (2021b) and Bianchi et al. (2021a) describe other examples of models that incorporate embeddings created by PLM (e.g. S-BERT, Reimers and Gurevych (2019)) to a traditional topic modeling structure, but without using a distance-based clustering algorithm to determine topics.

Models that directly cluster vectors in the vector space created by a PLM are what we call NTM-CE. The popularity of this approach increased after the release of BERT (Devlin et al., 2019), and most models since then have used BERT or a BERT variant. Sia et al. (2020) used BERT, principle component analysis (PCA) (Pearson, 1901; Hotelling, 1933), and K-Means clustering (Lloyd, 1982) and found the NTM-CE pipeline to perform similarly to LDA. Thompson and Mimno (2020) compared different flavors of BERT and GPT-2 (Radford et al., 2019) in combination with PCA and K-Means, also concluding that the technique performs better than LDA. Using more recent libraries, Grootendorst (2022) proposed BERTopic, which uses BERT, UMAP (McInnes et al., 2018), and HDBSCAN (Campello et al., 2013; McInnes and Healy, 2017) together with a novel term-weighting procedure c-TF-IDF to discover topics in news data. Similarly, Zhang et al. (2022) use BERT, UMAP, and K-Means together with term weighting and conclude that the model outperforms all previous models. While all of these models use PLM embeddings, there is no consensus as to which embeddings to cluster. Sia et al. (2020) clusters vocabulary-level embeddings, Thompson and Mimno (2020) clusters token-level embeddings and Grootendorst (2022) clusters sentence-level embeddings. Part of the current proposal is to investigate differences in properties between different embedding choices and to study whether one of them is preferable for NTM-CE.

Taking another perspective, Meng et al. (2022) jointly train the dimension reduction and clustering components to obtain a vector space with high clusterability in the sense of (Ackerman and Ben-David,

2009). Their model TopClus was highly successful and thus an interesting testament to what can be done when combining components. However, to limit the research plan described here, we will initially focus on approaches with separately trained components, saving jointly trained ones for later.

3 Proposed work

In this section, we present the proposed work to seek answers to the research questions posed in Section 1. The research is divided into four phases, expected to correspond to 1–2 papers each. The work is to be done over two years. Before explaining the phases, an outline of a potential system that is supposed to accompany the theoretical work is given.

3.1 Topic Modeling System

The preliminary framework for a dynamic NTM-CE is shown in Figure 1. As previously mentioned, the core of an NTM-CE model is *vectorization* \rightarrow *transformation* \rightarrow *clustering*. To make it a dynamic topic model, the component *temporalization* is added. The temporal functionality is loosely defined as the part of the system that adds the dynamic aspect to the topic model. For the discrete case, this could be the binning of the documents depending on their timestamps. For the continuous case, this could be relating the documents to a time function. Deciding how the temporalization should be designed and where it should be placed in the system is a core part of the research to be done.

In terms of limitations, systems for topic modeling usually have an additional component that roughly describes each of the topics in a human-readable format. This description usually takes the form of keywords. However, the exploration of topic descriptions is considered to be outside the scope of this research proposal as the aim of the project is to study the dynamic aspect. Another limitation of NTM-CE is that the topics need to be created by a distance-based clustering algorithm. This means that the system will have freedom in how to create and manipulate vectors, but that the vectors in the end must be susceptible to mentioned clustering algorithms.

3.2 Phases

Phase I: An intuitive NTM-CE model in discrete time that can be compared with other DTMs will be implemented. The previously mentioned experi-

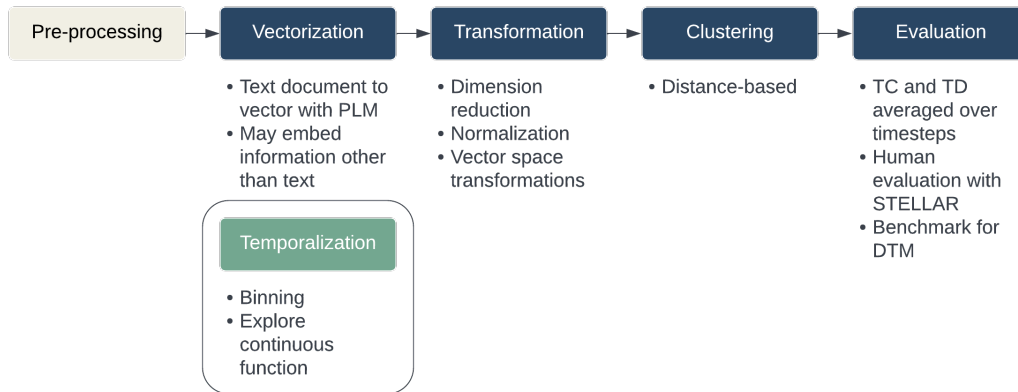


Figure 1: A preliminary framework in which to build dynamic topic models. The temporalization component is loosely defined as the part that adds functionality for making the topic modeling system dynamic. This includes binning documents to different timesteps and also a more explorative continuous function for NTM-CE. TC=Topic Coherence. TD=Topic Diversity.

ments by Grootendorst (2022) compared a binned BERTopic with d-LDA on the averaged topic coherence and topic diversity over a number of timesteps. However, we want to compare more models binned in this way, as well as models such as TOT and c-DTM. This work will establish a baseline for the comparison of dynamic topic models. Phase I will include the work to develop a code base for the comparison and evaluation of dynamic topic models. The evaluation, discussed further in Section 4, will combine automatic quantitative metrics and human evaluation. Moreover, familiarization with the strengths and weaknesses of existing models is crucial for continued research.

Phase II: An exploration of general NTM-CE properties is needed to work towards RQ1. Properties are often revealed in the vector space as patterns or structures that could be exploited. The requirements an ideal dynamic topic modeling would put on a system will be used to guide which properties to look for and to extract from NTM-CE. Properties could originate from any of the components described in Section 3.1, or a combination of them. A topic model based on the properties found will be developed. This model would likely compete with the state-of-the-art models or showcase some other features that are unique to dynamic NTM-CE.

Phase III: After developing a model in Phase II, the model will be generalized into a framework that is robust to future changes in components. A natural start is to generalize the framework to include Word2Vec-based models. With that, we can

see if the properties found for transformations are general enough for different PLMs or if we need to reconsider. The desired outcome of this phase is a framework that not only works for the specific components available today, such as current transformer-based models, but would allow replacing components in the pipeline with future, more advanced ones.

Phase IV: The last step after developing a framework will be to expand the evaluation and study further application areas. The project has an industry partner and will therefore have the unique opportunity to perform real-world evaluations on industry datasets of news articles, considering applications more relevant to those outside of academia. A planned study is to look at the news cycle spanning over at least two years to analyze events that reoccur, and events that emerge and then disappear. Another ongoing related project at our home university looks at the detection of formal narrative structures in news articles with the aim to use it in longitudinal studies of reporting. Moreover, our home university does extensive work in gender studies which opens up for similar studies around gender bias. For example, the dynamic topic model can be used to identify changes in the use of stereotypical gender roles in language, and, in extension, that understanding may help the debiasing of datasets used in NLP.

4 Evaluation

Evaluation of topic models is not trivial as the lack of an objective ground truth makes it hard

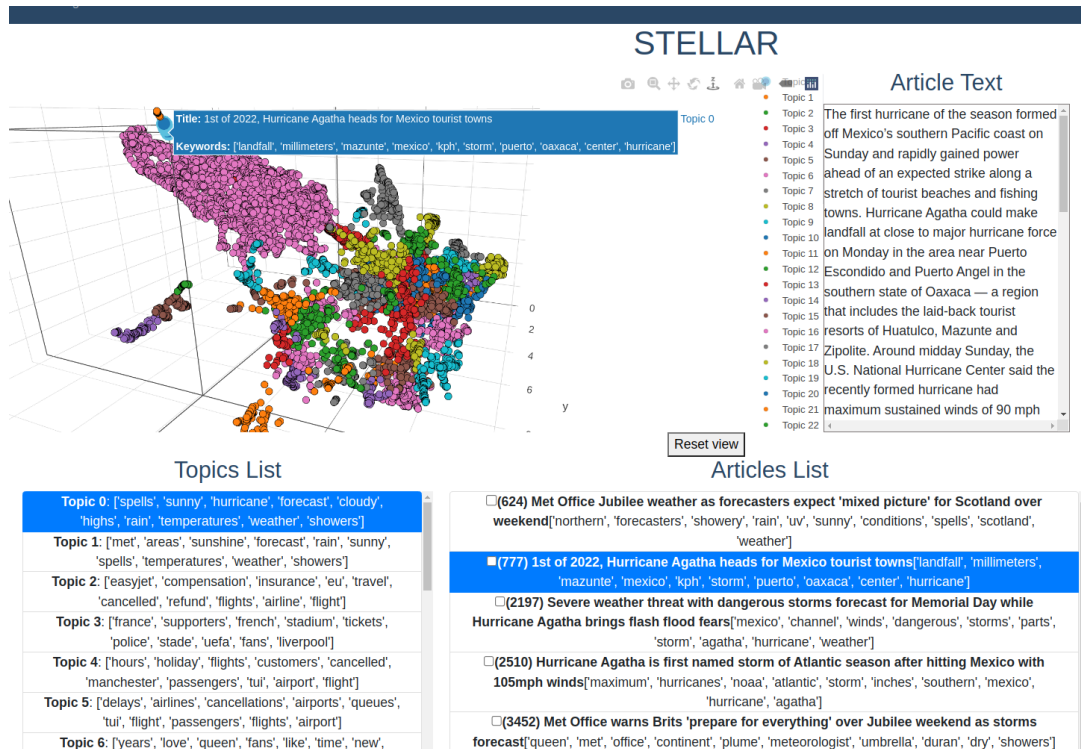


Figure 2: The topic browser that can be used for quick human evaluation of topics. The main functions are the 3D graph which shows a reduced version of the vector space, the list of topics with keywords, the list of articles in a topic, and the article text of a selected article.

to achieve a consensus on the number and nature of topics. Automatic measurements for evaluating topic models are topic coherence and topic diversity. While there are many approaches, topic coherence is usually automatically measured using normalized point-wise mutual information (NPMI, Bouma (2009)) and is considered by some to mimic human judgment (Lau et al., 2014). Topic diversity measures in different ways how diverse the top words in a topic are to each other (Bischof and Airoldi, 2012; Dieng et al., 2020; Bianchi et al., 2021a). There are tools like OCTIS (Terragni et al., 2021) that make it easy to compare static topic models. We plan to extend OCTIS or use similar ideas to facilitate fair comparisons between DTMs. As a first step, an OCTIS extension may average topic coherence and topic diversity over the time steps as has been done by Grootendorst (2022) and Dieng et al. (2019). Furthermore, part of the work in Phase I will be to assess how to measure aspects more specific to dynamic modeling requirements. This could, e.g., result in the requirement to develop an initial benchmark dataset for topic change over time, or to find a way to quantitatively assess topic change without a ground truth.

The importance of automatic measurements that

correlate with human judgment has been known (Chang et al., 2009) and NPMI was adopted after showing such a correlation. However, a recent study by Hoyle et al. (2021) argues that automatic coherence measurements, including the prevailing topic model evaluation standard NPMI, should not be considered equivalent to human judgment. Therefore, we plan to complement automatic measurements with human evaluation when resource allocation is justified, for example, when a core pillar of the work needs to be validated. Qualitatively, we assume that a human can look at a sample of topics produced by a topic model and decide if they think the topics are coherent and also which documents should not be considered to belong to the topic. We use this assumption to develop a tool for rapid human evaluation of topics, which we intend to make use of to validate automatic measurements. This tool is described further in Section 5

5 Preliminary Work

STELLAR² (Systematic Topic Evaluation Leveraging Lists of ARticles) (Eklund and Forsman, 2022) was developed as a tool for rapid human

²<https://github.com/antoneklund/STELLAR>

evaluation; see Figure 2. The idea behind it is that the coherence of a topic can be more confidently assessed if an evaluator reads the actual titles and text of the articles rather than only a few describing keywords. By having the evaluator systematically go through all topics, we can get a score for how well the model performs from a human perspective. This requires expert evaluators who can contextualize a given number of articles and then also mark articles that do not belong to the topic.

STELLAR is supposed to aid the evaluation process and make it faster. The core functionalities are a topic list, an article list from the chosen topic, a box for reading the article text body, and a 3D visualization of the document vector space. In the article list, the articles can be marked as not belonging to the topic. The tool will be extended to make it easier to analyze dynamic topic models. The extension could add functionality like selecting articles within different time periods, visualizing the varying size of topics over time, or visualizing changes in the topic description over time. The functions and statistics that are needed we expect to become clear over the course of working with this project.

6 Impact

In the current information era, there are obvious benefits to having fast and trusted topic models which can process large corpora of documents. We have seen that there is a wide range of applications such as analyzing historical documents, social media, or news articles. Research on NTM-CE is particularly interesting because of the modularity of the different components, especially to isolate the language model in the system. This allows for exchanging parts of the system when new and better components are developed, meaning that this type of topic modeling will continue to improve even after the popular language model at the time is superseded by a better one. Developing a solid framework for dynamic topic modeling with this modularity will ensure that NTM-CE models are as flexible as LDA-based models in their applications.

7 Summary

This paper presented an outline to explore the dynamic topic modeling in detail with NTM-CE. We proposed two research questions in an attempt to cover both the theoretical framework and practical application of dynamic topic modeling with NTM-

CE. We propose four phases in which the work will be done and where the main contributions will be: 1) a codebase for evaluating dynamic topic models, 2) a general framework for how to efficiently create dynamic topic models with NTM-CE, and 3) insights from practical application of the framework to various datasets. The research questions and phases were developed to the best of our ability with our current understanding. However, we see them as constantly evolving as we learn more. Therefore, we highly welcome all types of input from the research community to make this project as relevant and impactful as it can be.

Acknowledgements

This Ph.D. student is funded by the Swedish Foundation for Strategic Research, project id ID19-0055.

References

- Margareta Ackerman and Shai Ben-David. 2009. Clustering: A theoretical study. In *Artificial intelligence and statistics*, pages 1–8. PMLR.
- Federico Bianchi, Silvia Terragni, and Dirk Hovy. 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)*, pages 759–766. Online. Association for Computational Linguistics.
- Federico Bianchi, Silvia Terragni, Dirk Hovy, Debora Nozza, and Elisabetta Fersini. 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*, pages 1676–1683. Online. Association for Computational Linguistics.
- Jonathan Bischof and Edoardo M Airolidi. 2012. Summarizing topical content with word frequency and exclusivity. In *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, pages 201–208.
- David M. Blei and John D. Lafferty. 2006. Dynamic topic models. In *ICML '06: Proceedings of the 23rd international conference on Machine learning*, pages 113–120.
- David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022.
- Gerlof Bouma. 2009. Normalized (pointwise) mutual information in collocation extraction. *Proceedings of the Biennial GSCL Conference*.

- Ricardo J. G. B. Campello, Davoud Moulavi, and Joerg Sander. 2013. Density-based clustering based on hierarchical density estimates. In *Advances in Knowledge Discovery and Data Mining*, pages 160–172, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Mario Cataldi, Luigi Di Caro, and Claudio Schifanella. 2010. Emerging topic detection on twitter based on temporal and social terms evaluation. In *Proceedings of the tenth international workshop on multimedia data mining*, pages 1–10.
- Jonathan Chang, Sean Gerrish, Chong Wang, Jordan Boyd-Graber, and David Blei. 2009. Reading tea leaves: How humans interpret topic models. *Advances in neural information processing systems*, 22.
- Rob Churchill and Lisa Singh. 2021. [The evolution of topic modeling](#). *ACM Comput. Surv.* Just Accepted.
- Hannah Devinney, Jenny Björklund, and Henrik Björklund. 2020. [Semi-supervised topic modeling for gender bias discovery in English and Swedish](#). In *Proceedings of the Second Workshop on Gender Bias in Natural Language Processing*, pages 79–92, Barcelona, Spain (Online). Association for Computational Linguistics.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. [BERT: Pre-training of deep bidirectional transformers for language understanding](#). In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Adji B. Dieng, Francisco J. R. Ruiz, and David M. Blei. 2019. [The dynamic embedded topic model](#). *arXiv preprint arXiv:1907.05545*.
- Adji B. Dieng, Francisco J. R. Ruiz, and David M. Blei. 2020. [Topic modeling in embedding spaces](#). *Transactions of the Association for Computational Linguistics*, 8:439–453.
- Anton Eklund and Mona Forsman. 2022. Topic modeling by clustering language model embeddings: Human validation on an industry dataset. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: Industry Papers*. Just accepted.
- Maarten Grootendorst. 2022. [BERTopic: Neural topic modeling with a class-based TF-IDF procedure](#). *arXiv preprint arXiv:2203.05794*.
- Harold Hotelling. 1933. [Analysis of a complex of statistical variables into principal components](#). *Journal of Educational Psychology*, 24.
- Alexander Hoyle, Pranav Goel, Andrew Hian-Cheong, Denis Peskov, Jordan Boyd-Graber, and Philip Resnik. 2021. [Is automated topic model evaluation broken? The incoherence of coherence](#). In *Advances in Neural Information Processing Systems*, volume 34, pages 2018–2033. Curran Associates, Inc.
- Jey Han Lau, David Newman, and Timothy Baldwin. 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*, pages 530–539, Gothenburg, Sweden. Association for Computational Linguistics.
- Stuart P. Lloyd. 1982. [Least squares quantization in PCM](#). *IEEE Transactions on Information Theory*, 28(2):129–137.
- Leland McInnes and John Healy. 2017. [Accelerated hierarchical density based clustering](#). In *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE.
- Leland McInnes, John Healy, Nathaniel Saul, and Lukas Großberger. 2018. [Umap: Uniform manifold approximation and projection](#). *Journal of Open Source Software*, 3(29):861.
- Yu Meng, Yunyi Zhang, Jiaxin Huang, Yu Zhang, and Jiawei Han. 2022. [Topic discovery via latent space clustering of pretrained language model representations](#). In *Proceedings of the ACM Web Conference 2022, WWW '22*, page 3143–3152, New York, NY, USA. Association for Computing Machinery.
- Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. [Efficient estimation of word representations in vector space](#). *arXiv preprint arXiv:1301.3781*.
- David J Newman and Sharon Block. 2006. Probabilistic topic decomposition of an eighteenth-century american newspaper. *Journal of the American Society for Information Science and Technology*, 57(jun):753–767.
- Karl Pearson. 1901. [LIII. On lines and planes of closest fit to systems of points in space](#). *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2.
- Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.
- Nils Reimers and Iryna Gurevych. 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)*, pages 3982–3992, Hong Kong, China. Association for Computational Linguistics.

- Suzanna Sia, Ayush Dalmia, and Sabrina J. Mielke. 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)*, pages 1728–1736, Online. Association for Computational Linguistics.
- Silvia Terragni, Elisabetta Fersini, Bruno Giovanni Galuzzi, Pietro Tropeano, and Antonio Candelieri. 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*, pages 263–270. Association for Computational Linguistics.
- Laure Thompson and David Mimno. 2020. [Topic modeling with contextualized word representation clusters.](#) *arXiv preprint arXiv:2010.12626*.
- Chong Wang, David Blei, and David Heckerman. 2008. Continuous time dynamic topic models. In *Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence, UAI'08*, page 579–586, Arlington, Virginia, USA. AUAI Press.
- Xuerui Wang and Andrew McCallum. 2006. [Topics over time: A non-markov continuous-time model of topical trends.](#) In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '06*, page 424–433, New York, NY, USA. Association for Computing Machinery.
- Zihan Zhang, Meng Fang, Ling Chen, and Mohammad Reza Namazi Rad. 2022. [Is neural topic modelling better than clustering? An empirical study on clustering with contextual embeddings for topics.](#) In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 3886–3893, Seattle, United States. Association for Computational Linguistics.
- He Zhao, Dinh Phung, Viet Huynh, Yuan Jin, Lan Du, and Wray Buntine. 2021. [Topic modelling meets deep neural networks: A survey.](#) *arXiv preprint arXiv:2103.00498*.