JOIN (Jurnal Online Informatika)
p-ISSN: 2528-1682, e-ISSN: 2527-9165
Volume 8 Number 2 | December 2023: 147-157
DOI: 10.15575/join.v8i2.963

Implementation of Dynamic Topic Modeling to Discover Topic Evolution on Customer Reviews

Valentinus Roby Hananto
Department of Information Systems, Universitas Dinamika, Surabaya, Indonesia

Article Info

Article history:

Received October 31, 2022
Revised May 12, 2023
Accepted May 19, 2023
Published December 28, 2023

Keywords:

Dynamic topic modeling
BERTopic
Topic evolution
Customer reviews

ABSTRACT

Annotation and analysis of online customer reviews were identified as significant problems in various domains, including business intelligence, marketing, and e-governance. In the last decade, various approaches based on topic modeling have been developed to solve this problem. The known solutions, however, often only work well on content with static topics. As a result, it is challenging to analyze customer reviews that include dynamic and constantly expanding collections of short and noisy texts. A method was proposed to handle such dynamic content. The proposed system applied a dynamic topic model using BERTopic to monitor topics and word evolution over time. It would help decide when the topic model needs to be retrained to capture emerging topics. Several experiments were conducted to test the practicality and effectiveness of the proposed framework. It demonstrated how a dynamic topic model could handle the emergence of new and over-time-correlated topics in customer review data. As a result, improved performance was achieved compared to the baseline static topic model, with 25% of new segmented texts discovered using the dynamic topic model. Experimental results have, therefore, convincingly demonstrated that the proposed framework can be used in practice to develop automatic review annotation tools.

Corresponding Author:
Valentinus Roby Hananto,
Department of Information Systems, Faculty of Technology & Informatics, Universitas Dinamika
Jl. Raya Kedung Baru No.98, Rungkut, Surabaya, Indonesia. 60298
Email: valentinus@dinamika.ac.id

1. INTRODUCTION

The fast expansion of the Internet has increased the number of people who write evaluations on e-commerce platforms, resulting in an abundance of online reviews overall. From the user's standpoint, reading online product reviews from previous customers will assist them in better understanding the product's quality and functionalities [1]. Meanwhile, online reviews offer critical information for company shareholders that can help research and develop innovative products, improving customer services and quality control [2]. However, while websites with many informative reviews tend to attract more purchases [3], the sheer quantity of reviews of varying quality can be overwhelming for both consumers and businesses [4]. The vast number of consumer feedback available on e-commerce platforms presents a substantial challenge to organizations when collecting and analyzing relevant information about customer demands and expectations. If a product is top-rated, consumers are more likely to give a significant number of comments, which may include a variety of topics that could be derived, for example, from requirement elicitation [5]. This circumstance, often known as information overload [6], should be addressed so that consumer feedback can be utilized for product or service improvement [7]. A document annotation system is often proposed to solve information overload problems by selecting and categorizing reviews that have been recorded [8]–[11].

Topic modeling techniques, such as Latent Dirichlet Allocation (LDA), are unsupervised machine learning techniques that extract latent themes from text documents like online reviews. LDA methods have been used increasingly as an indispensable business intelligence tool to analyze various online documents [12], [13]. The successful application of topic modeling in solving these and many other semantic classification problems made it a popular approach to assist or even fully automate the data annotation process [14]. However, it is also important to note that dealing with dynamic, constantly flowing volumes of customer review data would require a method capable of handling dynamic content. Developing annotation tools for customer reviews would require the replacement of the basic topic modeling in the system with dynamic topic modeling, e.g., as proposed in [15]–[17]. Dynamic topic modeling, also known as DTM, is used to represent the evolution over time of prevalent themes in various kinds of text, including literature, online content, and customer reviews [16].

In addition to the topic modeling approach, text segmentation (TS) approaches have been widely used to identify boundaries of semantic units, such as sentences, paragraphs, and news, to mine opinions and emotions, evaluate sentiment, determine language, and various other applications [18]. Using text segmentation techniques, a document can be broken into smaller coherent units, usually called text segments. Text segmentation is essential in dealing with information overload problems that could happen when users are presented with a lengthy document.

Several studies have been conducted to implement Dynamic Topic Modeling (DTM) on customer reviews. Sousa and Becker [15] utilized DTM to summarize the topics expressed by Twitter users in Brazil regarding COVID-19 vaccination. Another study also conducted an analysis of evolving COVID-19 vaccine-related opinions [19]. Meanwhile, Tomasi *et al.* [17] identified the evolution of topics in the machine learning literature using DTM. A similar approach was proposed to discover the evolution of artificial intelligence in cancer studies [20]. Implementation of DTM was also done on the social media platforms [21].

This study proposes a system to handle dynamic content often found in online customer reviews. The proposed system applied a dynamic topic model to monitor topics and word evolution over time. It would help decide when the topic model needs to be retrained to capture emerging topics. As a result, improved performance was achieved compared to the baseline static topic model. In the previous study, an automatic annotation method was proposed using a text segmentation approach, but it still has limitations when dealing with topic evolution [22]. The main contribution of this study is a method that combines the application of dynamic topic modeling and text segmentation in order to discover topic evolution on customer reviews.

The rest of this paper is structured as follows. Section 2 introduces the suggested methods for automatic review annotation based on dynamic topic modeling. Several experiments implementing dynamic topic model to handle the emergence of new and over-time-correlated topics in customer reviews are described in Section 3. The results obtained from the experiments are then discussed. Finally, Section 4 formulates the conclusions and outlines the study's limitations.

2. METHOD

The proposed method starts with the data collector module, which prepares the required data format, followed by the topic model development module.

2.1. Data Collection and Pre-processing

The data collector crawls massive amount of customer review data from popular e-commerce platforms, for instance, from the Amazon dataset [23] that contains 5,789,920 reviews written in English between 1996 and 2014. It was necessary to gather relevant product IDs from inside the reviews' information in order to be able to extract reviews that were tied to specific product categories.

Text pre-processing aims to clean up the dataset by removing unnecessary information and minimizing noise. Essential text pre-processing tasks that are incorporated into the system are listed below:

1. Text cleaning is a task to delete irrelevant elements (such as numbers, stop words, symbols, and pronouns), duplicate data, and too short documents.
2. Tokenization is a task to slice documents into tokens (i.e., words).
3. Part of Speech (POS) filtering is a task to tag every word with each corresponding POS and then remove unneeded categories (e.g., interjections).

4. Lemmatization aims at reducing inflectional and derivational forms of a word to its standard dictionary form called lemma, based on morphological analysis and a dictionary (for example, all word "visitor," "visiting," and "visits" becomes "visit").
5. The N-gram model is used to search for multiword concepts (e.g., "swimming pool" and "ice cream").
6. The bag-of-words (BoW) model stores the pre-processed tokens in vectors that describe the occurrences of tokens within documents. The BoW model is required for training purposes of an LDA model.

2.2. Topic Model Development

The topic model development module focuses on converting the pre-processed text documents into a topic model. The topic modeling process is carried out to generate document clusters, which are subsequently used as a reference model in the text segmentation phase.

A. LDA Topic Modeling

LDA topic modeling works to cluster the pre-processed documents unsupervised, assuming that documents with similar content should be grouped regardless of their structure. The LDA algorithm is applied to the BoW model in an unsupervised fashion. The LDA application assumes that reviews with the same theme have content that is also similar, independent of the word structure. For every document (d), a distribution of k latent topics is computed that serves as the multinomial distribution over each word within the corpus.

One challenge in unsupervised clusterings, such as LDA, is deciding the optimal number of clusters. There are several techniques to evaluate a topic model. The C_v coherence score is one of the measurements to find the best number of topics in LDA [24]. The C_v coherence measurement can outperform all other configurations of coherence measures. Considering $W = \{w_1, ..., w_N\}$ as the top-N words of a topic, a segmentation S is a subset of pairs from W, denoted as $S = (W', W^*), W', W^* \subseteq W$. W^* as the second part of the pair is used to confirm the first part W'. Common words that have high frequency are more likely to appear on documents. Remarkably, the domain words co-occurring strongly in the real world are scarce. Normalized Pointwise Mutual Information (NPMI) adjusts the frequencies to make them more easily interpretable and less sensitive to high-frequency, co-occurring words. Each word w_i is represented by an N-dimensional vector $\vec{v}(w_i) = \{NPMI\}_{j=1,\dots,N}$ where j-th entry is the NPMI between word w_i and w_i. It is calculated as follows:

$$NPMI = \frac{log_2 P(w_i, w_j) - log_2 (P(w_i), P(w_j))}{-log_2 (P(w_i, w_j) + \epsilon)} \tag{1}$$

where P is the probability of words and ϵ is a small constant added to prevent a logarithm of zero. After that, ϕS_i the confirmation measure of a pair S_i is then generated by computing the cosine vector similarity of all context vectors (with $\vec{u} = \vec{v}(W')$ and $\vec{w} = \vec{v}(W^*)$) by using their dot product and magnitude:

$$\phi S_i(\vec{u}, \vec{w}) = \frac{\vec{u}.\vec{w}}{||\vec{u}|| ||\vec{w}||} \tag{2}$$

The final C_v score is an aggregate of the ϕS_i for each word group pair into a single score based on the arithmetic mean of the given topic:

$$C_v = \frac{1}{N} \sum_{i=1}^{N} \phi S_i(\vec{u}, \vec{w}) \tag{3}$$

Utilizing this score allows for comparing topic models derived from a broad range of topic numbers. C_v calculations perform counterintuitively when the corpus contains many but irrelevant words. Since it occurs commonly in many corpora, where a standard vocabulary is often repeated in the textual data but is generally uninformative, pre-processing is necessary to alleviate this issue.

B. Dynamic Topic Modeling

Traditional topic modeling approaches are naturally static and thus do not allow for the modeling of documents ordered sequentially. It is understood that dealing with dynamic, constantly flowing volumes of customer review data would require transitioning to dynamic topic modeling.

BERTopic, proposed by Grootendorst [25], has progressively gained popularity as a tool for dynamic topic modeling [15], [26], [27]. Figure 1 shows the main components of the BERTopic algorithm.

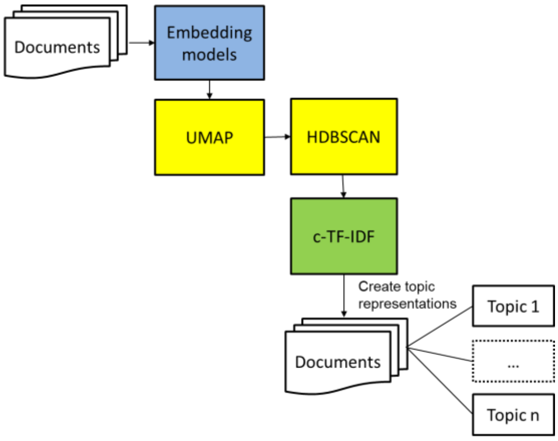

Figure 1. BERTopic main components

Four critical components used in BERTopic [25] are as follows:

1. A transformer embedding model. BERTopic supports several libraries for encoding the text into dense vector embeddings. The Sentence Transformers library provides the most extensive high-performing sentence embedding models.
2. The Uniform Manifold Approximation and Projection (UMAP) [28] is utilized for the dimensionality reduction of the embedding model. Following the construction of the embedding, BERTopic will compress them into a space with fewer dimensions. The high-dimensional vectors are transformed into two/three-dimensional vectors because high-dimensional vectors to represent text are unlikely to be needed. By applying UMAP, the following clustering step can be done more efficiently.
3. Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) is a hierarchical and density-based clustering method. The method benefits from the easier tuning and visualization of hierarchical data, the handling of irregular cluster shapes, and the identification of outliers [29].
4. Cluster tagging using c-TF-IDF is the final step in BERTopic to extract topic representation for each cluster. c-TF-IDF is a modified variant of the term frequency-inverse document frequency (TF-IDF) that BERTopic employs for this purpose.

To interpret topics, BERTopic proposes to associate the terms with the measure c-TF-IDF, which calculates the TF-IDF of the terms by class. The following formula is used to determine the c-TF-IDF scores for each class i:

$$c - TF - IDF = \frac{t_i}{w_i}\, x \log \frac{m}{\sum_{j=1}^{n} t_j} \tag{4}$$

where t_i is the word t frequency in class i, w_i is the number of words in class i, m is the number of documents, and $\sum_{j=1}^{n} t_j$ refers to the total frequency of word t across all classes n. Words with very high c-TF-IDF scores can be selected for each topic to evaluate the most important terms of each topic and their relation to the others (for example, in a specific period). By adding a temporal dimension to the global representation of the main topics extracted, BERTopic allows us to create a distribution and visualize how topics are represented at different points in time. With the temporal dimension, the representation of the topic at the moment may differ from the global representation. Users can define when the referenced topic model should be updated by monitoring this temporal dimension of BERTopic.

C. Text Segmentation

In this study, a text segmentation method called TopicDiff-LDA is employed to segment text automatically. The method works as illustrated in the flowchart below (see Figure 2).

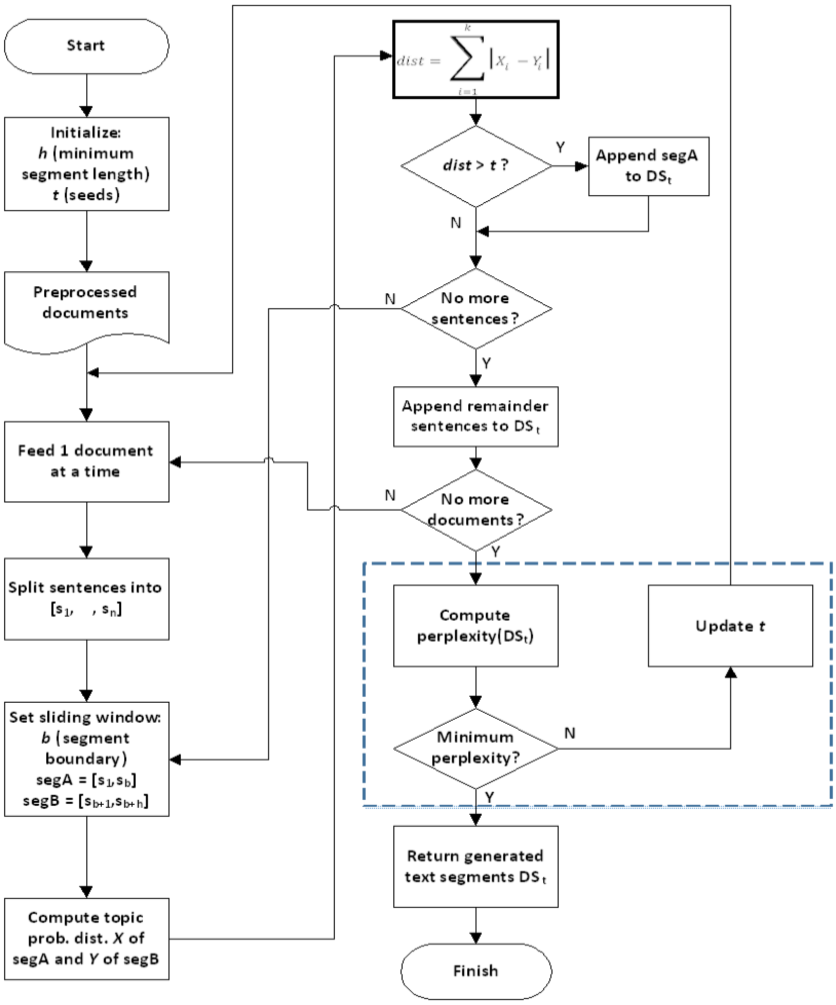

Figure 2. Flowchart of text segmentation method

The number of topics covered by the trained LDA model corresponds to the size of the vector, denoted by k, that is the optimal number of topics determined by the C_v score. To optimize the model, the objective function is to minimize $Perplexity$, which is a measure of assessing how good a segment model is [30]. Better models have lower perplexity, indicating a lower uncertainty about the text. Higher word likelihood is expected when words often occur in a segment, resulting in a lower perplexity. $perplexity_{N_S}$ is the perplexity score for N_S segments in the document collection, corresponding to a given distance threshold t. It is computed as follows:

$$perplexity_{N_S} = exp\left\{-\frac{\sum_{i=1}^{N_S}\log p(w_i)}{\sum_{i=1}^{N_S}N_{w_i}}\right\} \quad (5)$$

where N_{w_i} represents the number of words in segment i, and $\log p(w_i)$ is the log-likelihood of words in the segment.

3. RESULTS AND DISCUSSION

An experiment was designed on customer reviews to help users analyze the topicality change over time. This experiment conducts a temporal study to understand the behavior of Amazon customers when purchasing digital cameras.

3.1. Data

Customer reviews used in this research come from the "digital camera" part of Amazon review data [23]. An example of a review on the Amazon website is shown in Figure 3. From the dataset, 236,975 reviews were retrieved using the extracted IDs, and the number of unique products was 7,673. Reviews with less than 20 tokens were presumed too short and not meaningful; therefore, these were excluded from further processing steps.

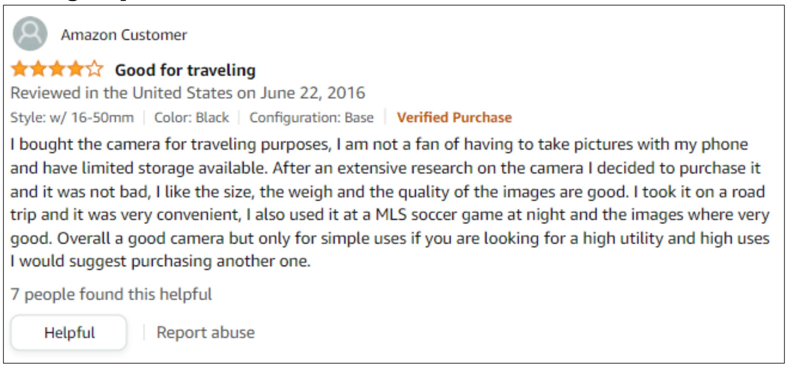

Figure 3. An example of a digital camera review from Amazon

3.2. Static Topic Modeling

To understand whether topics underlying each stance change over time, two static topic models were trained on two different periods of the collected data. This way, it is assumed that their representation of topics would be different. One possible way to divide the dataset is based on the review sequential order. Thus, the dataset was split into set A, the first 35,000 reviews (from July 1999 to December 2007), and set B, the last 35,000 reviews (from December 2015 to September 2018). Table 1 outlines the dataset structure. Deciding on the number of clusters for topic modeling was addressed by computing the C_v coherence scores to detect the optimal number of topics (k). The optimal number of topics was defined at k = 40 for set A and k = 50 for set B, as seen in Figure 4. The model generation for each of the two sets was done independently. Therefore, the obtained models do not affect each other's output.

Table 1. Dataset after splitting to set A for the first 35,000 reviews and set B for the last 35,000 reviews

Product	Set A	Set B
Number of unique tokens	75,842	39,715
Average words per review	170.87	91.48
Average sentences per review	8.37	5.31
Total number of segments	63,444	45,513

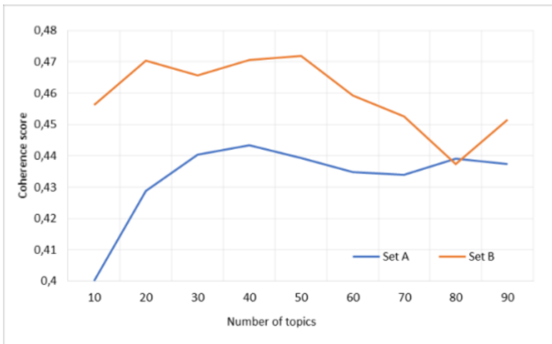

Figure 4. Change of C_v over the number of topics

Table 2 shows the most representative terms of the main topics extracted for both datasets. Topics were sorted by their number of reviews, and the top 11 were selected for comparison. The expressions "image," "quality," "feature," and "resolution" that exist in the Topic 1 on both sets represent the general topic of digital camera purchase. The rest of the topics have other similarities but also differences. In topic 2 of set A, the terms "problem" and "repair" show negative sentiment toward the topic.

In contrast, in set B, the terms "return" and "issue" in topic 5 indicate similar concerns about digital cameras. Topic 9 (waterproof camera) is a new topic that emerged in set B and did not exist in set A. The following experiments attempt to utilize dynamic topic modeling that is intended to capture any new over-time-correlated topics (e.g., waterproof camera).

Table 2. Generated top words on each topic

Topic	#reviews	Top words
Set A		
Topic	**#reviews**	**Top words**
1	2,489	digital, image, quality, feature, excellent, resolution
2	1,644	problem, repair, buy, send, fix, warranty
3	1,407	battery, life, charge, rechargeable, last, power
4	1,333	small, size, pocket, carry, fit, little, case
5	1,323	flash, light, low, shot, indoor, focus, dark
6	1,181	review, read, manual, buy, think, know
7	1,069	price, buy, quality, money, well, spend, worth
8	735	screen, lcd, see, viewfinder, large, view
9	556	mode, setting, button, manual, menu, change
10	541	video, movie, sound, record, zoom, clip
11	512	software, computer, work, download, window, driver
Set B		
Topic	**#reviews**	**Top words**
1	2,660	love, quality, amazing, price, awesome, feature
2	2,411	photography, learn, love, want, start, easy, beginner
3	1,623	package, come, product, arrive, happy, receive
4	1,541	tripod, bundle, bag, accessory, extra, include
5	1,526	work, time, try, return, problem, issue
6	1,409	small, carry, pocket, size, easy, fit, travel, light
7	915	shot, light, focus, low_light, flash, setting
8	810	battery, charge, life, charger, last, extra
9	631	water, underwater, waterproof, beach, vacation
10	536	wifi, phone, transfer, app, connect, computer
11	534	mode, manual, setting, auto, shoot, control

3.3. Dynamic Topic Modeling

BERTopic was implemented as a dynamic modeling technique. In order to train the dynamic topic model, sentence embedding was implemented using all-mpnet-base-v2. This sentence-transformer model was pre-trained to map sentences and paragraphs to a dense vector space of 768 dimensions. In the experiment, the most influential parameters were *min_topic_size* (the minimum topic size) and *nr_topics* (the specified number of topics). *min_topic_size* was initially set to 500, resulting in 51 topics being generated. The hierarchical clustering of the dataset's 51 topics is presented in Figure 5. The more closely connected topics share more terms in common.

Some of the tools provided by the framework were used to cut the number of topics down to *nr_topics* because it is challenging to interpret and analyze such a vast number of topics. Calculating the c-TF-IDF matrix of the documents and then iteratively merging the least frequent topic with the most comparable one based on their respective c-TF-IDF matrices were the two steps to reduce the total number of topics. Following several iterations, the final total number of topics (*nr_topics)* was 15. The topics, as well as their sizes and representations, were then updated.

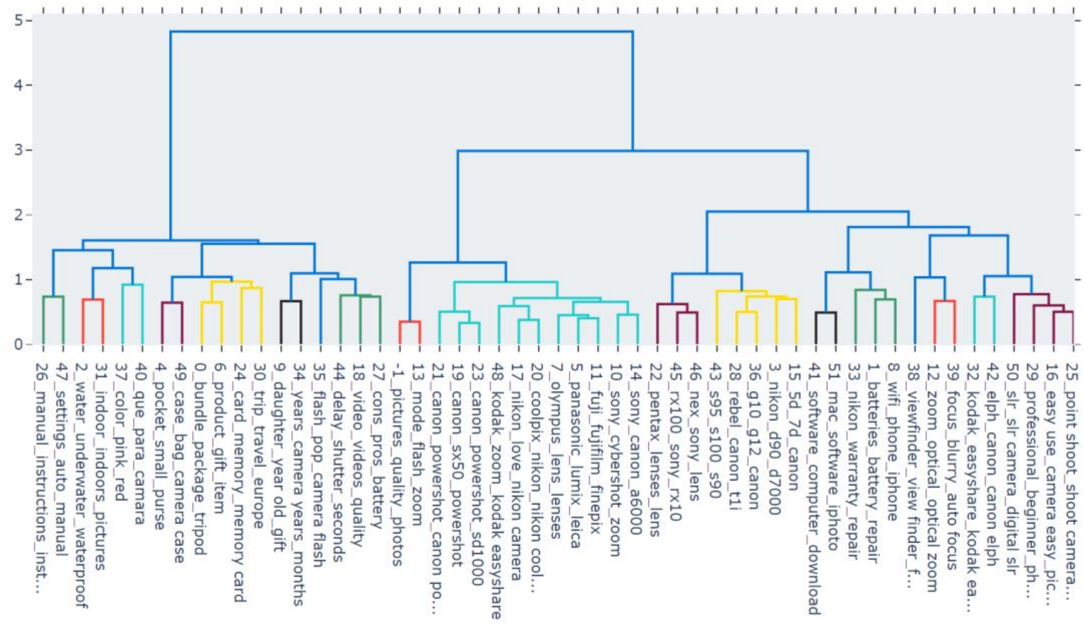

Figure 5. BERTopic hierarchical clustering

The BERTopic framework includes resources for viewing the topics found and the most important terms. It is also possible to arrange the volume of documents on each topic over time and extract significant terms at specific time intervals. Figure 6 shows the distribution of the topics described in Table 3 over time. In order to assess the evolution of topics over time, the topic with the highest volume of reviews in each position is described. Notably, starting in 2013, the normalized frequency of newer topics such as topics "waterproof" peaked. One would imply that 2013 is a recommended point to split the dataset and update the topic model. By monitoring the evolution of the topic "waterproof" in the chart, one can notice that the frequency has increased gradually since 2002, indicating a rising trend toward underwater digital cameras.

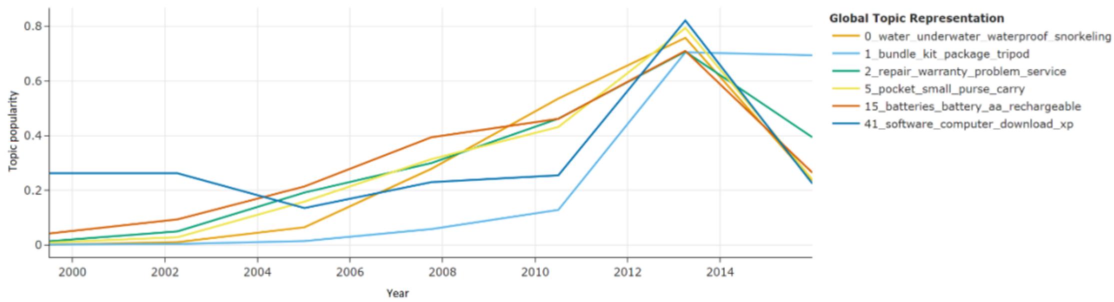

Figure 6. BERTopic monitoring topics over time

3.4. Discussion

Unlike the previous experiments, which only employed static topic modeling using LDA, dynamic topic modeling using the BERTopic was employed to recognize topic change over time. The temporal analysis included shows the evolution of the arguments represented by the topics over time. Based on the temporal topic suggested by the dynamic topic modeling, it was then decided to incorporate the newly seen topic and update the topic modeling. The updated topic model was then used to implement text segmentation of the TopicDiff-LDA, and the resulting segments were assessed using the C_v scores. Table 3 shows the C_v scores differences for each class before and after TopicDiff-LDA was implemented, with the growth of segment sizes around 25% on average. It is important to note that the new model performed significantly better ($\boldsymbol{p} < 0.001$) in all topic clusters.

Table 3. Changes in C_v coherence and segment growth after incorporating new topics in the text segmentation

Topic	Labels	Previous C_v	New C_v	Segment Growth (%)
1	general	0.377	0.461	11%
2	photography	0.386	0.465	21%
3	package	0.426	0.482	20%
4	accessories	0.493	0.549	15%
5	problems	0.509	0.567	14%
6	sizes	0.362	0.403	35%
7	light	0.446	0.477	32%
8	battery	0.489	0.538	32%
9	waterproof	0.479	0.535	41%
10	connectivity	0.509	0.613	28%
11	mode	0.514	0.655	28%

By manually inspecting the resulting segmented texts, it was found that there are several short reviews containing multiple topics. For example, a review contains the user's opinion about the camera's waterproof ability and also the experience when a part of the camera was broken. As can be seen below, this kind of review was challenging to segment and classify since it contains relatively new topics, such as waterproof cameras.

```
"I love this camera. I think I have had it for about two years. It is waterproof and shockproof. I have taken it razor riding (side by side) and it works well and nothing has happened to it. I will definitely buy this again! It is a great size. I am not sure what happened though. I took it out and the screen was broken. Recently I have been carrying it in my purse since we are redoing a house."
```

This review was initially clustered as "problems". After the new topic is incorporated following the implementation of dynamic topic model, the review can be split and the topic can be identified as "waterproof" and "problems", respectively:

`"I love this camera. I think I have had it for about two years. It is waterproof and shockproof. I have taken it razor riding (side by side) and it works well and nothing has happened to it."` (waterproof)

`"I will definitely buy this again! It is a great size. I am not sure what happened though. I took it out and the screen was broken. Recently I have been carrying it in my purse since we are redoing a house."` (problems)

This example illustrates that it was possible to identify the topic of "waterproof" after the model was retrained, using dynamic topic modeling. This topic, which appeared in the latter stage of the dataset, had the most significant segment size growth (41%) when the text segmentation was performed (see Table 3). This outcome, therefore, emphasizes the significant contribution of dynamic topic modeling in the proposed method.

4. CONCLUSIONS

Existing topic modeling methods are inherently static and hence cannot handle sequentially ordered documents that include new and over-time-correlated topics. Dynamic topic modeling allows for the incorporation of new, never-before-seen themes. To the best of our knowledge, there is limited application of dynamic topic modeling as a monitoring tool to capture topic evolution in classifier-based methods. The method proposed in this study, which integrates dynamic topic modeling and text segmentation, effectively worked to discover the topic evolution on customer reviews. The experiment demonstrated that 25% of new segmented texts with new topics were generated by dynamic topic modeling.

A part of the proposed method employs topic modeling as a generative technique for discovering co-occurrence in text documents. The benefit of such a generative model is that there is no language reliance since no external knowledge is involved. Hence, the suggested approach is potentially language-agnostic. However, this capability has not been further examined in the study. In future works, it is recommended to evaluate the established framework with languages other than English. It is also recommended to fully integrate dynamic topic modeling to automatically detect new topics and classify relevant documents in business intelligence applications with diverse languages.

5. REFERENCES

[1] M. Sun, "How Does the Variance of Product Ratings Matter?," *Manage. Sci.*, vol. 58, no. 4, pp. 696–707, Dec. 2011, doi: 10.1287/MNSC.1110.1458.

[2] N. Bashir, K. N. Papamichail, and K. Malik, "Use of Social Media Applications for Supporting New Product Development Processes in Multinational Corporations," *Technol. Forecast. Soc. Change*, vol. 120, pp. 176–183, Jul. 2017, doi: 10.1016/J.TECHFORE.2017.02.028.

[3] A. Qazi, K. B. Shah Syed, R. G. Raj, E. Cambria, M. Tahir, and D. Alghazzawi, "A concept-level approach to the analysis of online review helpfulness," *Comput. Human Behav.*, vol. 58, pp. 75–81, May 2016, doi: 10.1016/J.CHB.2015.12.028.

[4] G. O. Diaz and V. Ng, "Modeling and Prediction of Online Product Review Helpfulness: A Survey," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Jul. 2018, pp. 698–708. Accessed: May 06, 2022. [Online]. Available: https://www.cse.msu.edu/

[5] M. G. Parente, "Using NLP and Information Visualization to Analyze App Reviews," Master Thesis, Utrecht University, Utrecht, Netherlands, 2018. [Online]. Available: https://dspace.library.uu.nl/bitstream/handle/1874/368082/MScThesis_MGarciaParente.pdf?sequence=2

[6] Z. Jin, W. Zhangwen, and N. Naichen, "Helping consumers to overcome information overload with a diversified online review subset," *Front. Bus. Res. China*, vol. 13, no. 1, pp. 1–25, Dec. 2019, doi: 10.1186/S11782-019-0062-1/TABLES/8.

[7] H. Lee, K. Choi, D. Yoo, Y. Suh, S. Lee, and G. He, "Recommending valuable ideas in an open innovation community: A text mining approach to information overload problem," *Ind. Manag. Data Syst.*, vol. 118, no. 4, pp. 683–699, 2018, doi: 10.1108/IMDS-02-2017-0044/FULL/PDF.

[8] F. Pech, A. Martinez, H. Estrada, and Y. Hernandez, "Semantic Annotation of Unstructured Documents Using Concepts Similarity," *Sci. Program.*, vol. 2017, 2017, doi: 10.1155/2017/7831897.

[9] A. Canito, G. Marreiros, and J. M. Corchado, "Automatic Document Annotation with Data Mining Algorithms," *Adv. Intell. Syst. Comput.*, vol. 930, pp. 68–76, Apr. 2019, doi: 10.1007/978-3-030-16181-1_7.

[10] S. Tuarob, L. C. Pouchard, P. Mitra, and C. L. Giles, "A generalized topic modeling approach for automatic document annotation," *Int. J. Digit. Libr.*, vol. 16, no. 2, pp. 111–128, Mar. 2015, doi: 10.1007/S00799-015-0146-2/TABLES/3.

[11] K. Bontcheva and H. Cunningham, "Semantic Annotations and Retrieval: Manual, Semiautomatic, and Automatic Generation," in *Handbook of Semantic Web Technologies*, Springer Berlin Heidelberg, 2011, pp. 77–116. doi: 10.1007/978-3-540-92913-0_3.

[12] J. Qiang, Z. Qian, Y. Li, Y. Yuan, and X. Wu, "Short Text Topic Modeling Techniques, Applications, and Performance: A Survey," *IEEE Trans. Knowl. Data Eng.*, pp. 1–1, May 2020, doi: 10.1109/TKDE.2020.2992485.

[13] E. Gallinucci, M. Golfarelli, and S. Rizzi, "Advanced topic modeling for social business intelligence," *Inf. Syst.*, vol. 53, pp. 87–106, Oct. 2015, doi: 10.1016/J.IS.2015.04.005.

[14] X. Liao, Z. Zhao, X. Liao, and Z. Zhao, "Unsupervised Approaches for Textual Semantic Annotation, A Survey," *ACM Comput. Surv.*, vol. 52, no. 4, pp. 1–45, Aug. 2019, doi: 10.1145/3324473.

[15] A. M. de Sousa and K. Becker, "Pro/Anti-vaxxers in Brazil: a temporal analysis of COVID vaccination stance in Twitter," in *9th Symposium on Knowledge Discovery, Mining, and Learning, (KDMILE) 2021*, Oct. 2021, pp. 105–112. doi: 10.5753/KDMILE.2021.17467.

[16] Patrick Jähnichen, Florian Wenzel, Marius Kloft, and Stephan Mandt, "Scalable Generalized Dynamic Topic Models," in *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, 2018, pp. 84:1427-1435. Accessed: May 19, 2022. [Online]. Available: https://proceedings.mlr.press/v84/jahnichen18a.html

[17] Federico Tomasi, Praveen Chandar, Gal Levy-Fix, Mounia Lalmas-Roelleke, and Zhenwen Dai, "Stochastic Variational Inference for Dynamic Correlated Topic Models," in *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence*, 2020, pp. 124:859-868. Accessed: May 19, 2022. [Online]. Available: https://proceedings.mlr.press/v124/tomasi20a.html

[18] I. Pak and P. L. Teh, "Text Segmentation Techniques: A Critical Review," in *Innovative Computing, Optimization and Its Applications: Modelling and Simulations*, I. Zelinka, P. Vasant, V. H. Duy, and T. T. Dao, Eds. Cham: Springer International Publishing, 2018, pp. 167–181. doi: 10.1007/978-3-319-66984-7_10.

[19] A. Sharma, S. Susan, A. Bansal, and A. Choudhry, "Dynamic Topic Modeling of Covid-19 Vaccine-Related Tweets," *ACM Int. Conf. Proceeding Ser.*, pp. 79–84, Feb. 2022, doi: 10.1145/3528114.3528127.

[20] S. Mosallaie, M. Rad, A. Schiffauerova, and A. Ebadi, "Discovering the evolution of artificial intelligence in cancer research using dynamic topic modeling," *COLLNET J. Sci. Inf. Manag.*, vol. 15, no. 2, pp. 225–240, Jul. 2021, doi: 10.1080/09737766.2021.1958659.

[21] R. Churchill and L. Singh, "Dynamic Topic-Noise Models for Social Media," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13281 LNAI, pp. 429–443, 2022, doi: 10.1007/978-3-031-05936-0_34/COVER.

[22] V. R. Hananto, U. Serdült, and V. Kryssanov, "A Text Segmentation Approach for Automated Annotation of Online Customer Reviews, Based on Topic Modeling," *Appl. Sci.*, vol. 12, no. 7, p. 3412, Mar. 2022, doi: 10.3390/APP12073412.

[23] R. He and J. McAuley, "Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering," in *25th International World Wide Web Conference, WWW 2016*, Apr. 2016, pp. 507–517. doi: 10.1145/2872427.2883037.

[24] M. Röder, A. Both, and A. Hinneburg, "Exploring the space of topic coherence measures," in *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining - WSDM '15*, 2015, pp. 399–408. doi: 10.1145/2684822.2685324.

[25] M. Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," *arXiv:2203.05794 [cs.CL]*, Mar. 2022, doi: 10.48550/arxiv.2203.05794.

[26] N. F. F. d. Silva *et al.*, "Evaluating Topic Models in Portuguese Political Comments About Bills from Brazil's Chamber of Deputies," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13074 LNAI, pp. 104–120, 2021, doi: 10.1007/978-3-030-91699-2_8/FIGURES/4.

[27] C. Y. K. Williams, R. X. Li, M. Y. Luo, and M. Bance, "Exploring patient experiences and concerns in the online Cochlear Implant community: a natural language processing approach," *Clin. Otolaryngol.*, vol. 48, no. 3, pp. 442–450, Mar. 2023,

doi: https://doi.org/10.1111/coa.14037.
[28] L. McInnes, J. Healy, N. Saul, and L. Großberger, "UMAP: Uniform Manifold Approximation and Projection," *J. Open Source Softw.*, vol. 3, no. 29, p. 861, Sep. 2018, doi: 10.21105/JOSS.00861.
[29] L. McInnes, J. Healy, and S. Astels, "HDBSCAN: Hierarchical density-based clustering," *J. Open Source Softw.*, vol. 2, no. 11, p. 205, Mar. 2017, doi: 10.21105/JOSS.00205.
[30] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, 2003.