Available online at
https://journal.uitm.edu.my/ojs/index.php/JIKM/

e-ISSN: 2289-5337

Journal of Information and Knowledge Management (2025) Vol. 15 No. 1

Journal of Information and Knowledge Management

Dirichlet Multinomial Modelling Approaches in Analysing Anxiety Therapy Messages

Teh Faradilla Abdul Rahman[1]*, Norshita Mat Nayan[2], Nurhilyana Anuar[1], Aminatul Solehah Idris[1], Raudzatul Fathiyah Mohd Said[1]

[1]*Centre of Foundation Studies, Universiti Teknologi MARA, Cawangan Selangor, Kampus Dengkil, 43800, Selangor, Malaysia*
[2]*Institute of Visual Informatics, Universiti Kebangsaan Malaysia, Bangi 43600, Selangor, Malaysia*

ARTICLE INFO

Article history:
Received 01 August 2024
Revised 07 March 2025
Accepted 17 March 2025
Online first
Published 01 April 2025

Keywords:
Topic modelling
Text analysis
DMM
Anxiety topic

ABSTRACT

Despite the effectiveness of anxiety therapy through text messages, limited re-search was found to analyze the topics included in the therapy session. It is also unclear of which topic modelling approaches is the best in extracting anxiety therapy topics from text messages. Thus, this study aims to compare the performance of four topic modelling methods, namely Latent Feature Dirichlet Multinomial Mixture (LFDMM), Gibbs Sampling Dirichlet Multinomial Mixture, Generalized Polya-urn Dirichlet Multinomial Mixture and Poisson-based Dirichlet Multinomial Mixture Model on 28 text messages of anxiety-therapy. Four combinations of parameter settings were applied in the experiments to compare and decide the most suitable ones for future analysis. The performance of the topic modelling was evaluated using classification accuracy, clustering, and coherence scores. LFDMM has the best accuracy (34.10%) and clustering scores (0.5000, 0.4808) with combinations of hyperparameters $\alpha = 0.1$ and $\beta = 0.01$ to infer more relevant topics. This study contributes to the growing body of research on topics in anxiety therapy, offering insights into the role of topic modelling in shaping valuable content. The findings highlight the potential of topic modelling for therapy con-tent exploration, aiding in text messages strategies for anxiety intervention.

INTRODUCTION

Anxiety is one of the highly prevalent common mental disorders in the population. Anxiety affects a person's mood or feelings, while the strength and duration of their symptoms can range from mild to severe (from months to years). To date, in-person psychotherapy underpinning individual emotions and well-being. However, in-person psychotherapy has been significantly and specifically impacted by the COVID-19 pandemic. Thus, therapists are adjusting to primarily implementing remote therapy via mobile, text messaging. Recent research suggests that text message-based therapy with psychological theoretical

[1*]Corresponding author. *E-mail address*: tehfaradilla@uitm.edu.my

guidance is effective for reduction in anxiety level (Aguilera et al., 2021; Ritvo et al., 2021), improvement on treatment adherence and increase in treatment length. Although effective mental health text-based interventions have been developed, little is known about psychological topics that were sent by mental health professionals to patient in the treatment of anxiety. There could be hidden topics in text messages sent by mental health professionals targeted to anxiety and it is tiresome to find topics in a large corpus manually and this method is prone to mistakes. Consequently, investigating technological approaches to enhance latent topic search in huge collection of documents and overcoming the difficulty in select-ing good topic for patient becomes superior in the anxiety text messaging therapy.

This study aims to determine the performance of one of the technological approaches in text mining known as topic modelling methods in analyzing text messages for anxiety therapy and which of the topic modelling works best for this type of data. Specifically, this study addresses two research questions:

1. How do the performances of four different topic modeling methods compare in analyzing anxiety text messages?
2. Which topic modeling method is most effective for analyzing anxiety text messages?

The finding of this study will be insightful in understanding the efficiency of topic modelling in uncover latent text messages content, providing information into how to personalize a suitable topic for a respective anxiety patient individually.

LITERATURE REVIEW

Text Analysis using Topic Modelling

Text mining is a cutting-edge approach that enables researchers to process a large amount of textual data by segmenting and extracting the essential information. One of the text mining approaches is topic modelling methods which have been introduced by few scholars (Blei et al., 2003; Li et al., 2016; Nguyen et al., 2015; Yan et al., 2013; Zuo, Wu, et al., 2016; Zuo, Zhao, et al., 2016) and are proved to be power-ful tools in health research (Fairie et al., 2021), prediction of research trends (Gupta et al., 2022), business analysis (Kao & Luarn, 2020) and public responses to COVID-19 (Xie et al., 2021). The earliest topic modelling invented was Latent Dirichlet Allocation (LDA) which is used by most studies as the main topic modelling tool for analyzing long text (Voskergian et al., 2023). However, LDA is less efficient when working with short text. The disadvantage of LDA contributed to less reliable topic inference and resulted in less coherent topics across the short text analysis. Fortunately, researchers had come with other topic modelling algorithms to overcome this issue. This was done by customized and well-tuned some parameters for sparse documents (Voskergian et al., 2023). Dirichlet multinomial mixture (DMM) is one of the advanced topic modellings that worked well with short text (Voskergian et al., 2023). Unlike the LDA, DMM uses algorithm that assumes each document contains one hidden topic.

In one study, GSDMM and LDA were experimented on Tweets related to COIVD-19 (Weisser et al., 2023). The study implemented Pseudo-Document Simulation embedded with GSDMM and need a prior separate hyperparameter optimization which eventually resulted a better performance by GSDMM. A detailed exploration of the topics of text-based treatment has the ability to offer insightful information on particular issues, given the enormous volume of texts that mental health professionals sent to people with anxiety. To our knowledge, no recent topic modelling analysis has been reported on topics in text messages for the treatment of anxiety. Thus, this study aims to compare the performance of four topic modelling methods on anxiety-therapy text messages. This study experimented topic modelling that based on Dirichlet Multinomial Mixture (DMM) including Latent Feature Dirichlet Multinomial Mixture (LFDMM), Gibbs Sampling Dirichlet Multinomial Mixture (GSDMM), Generalized Polya urn Dirichlet Multinomial Mixture (GPUDMM) and Poisson-based Dirichlet Multinomial Mixture Model (GPU-PDMM). Findings from the mentioned scholars shows that each distinct topic modelling performs differently on different dataset.

Consequently, there is a need to compare and determine more efficient topic modelling that can aid in detecting and analyzing text messages content used by mental health professionals for anxiety therapy.

Dirichlet Multinomial Mixture-based Topic Modelling

DMM is a straightforward and effective model that has been used to infer latent topics in short texts. The DMM model assumes the document collection follows a topic distribution and each short text is generated with a single topic. Many studies have proven that the DMM model consistently performs well when handling short texts (Qiang et al., 2019). However, this model's limitation will impact the general level of topic reasoning quality. Thus, the simple and effective DMM model has been extended by scholars to improve the performance in finding latent topics in variety of short text-related tasks. Yin & Wang (2014) proposed GSDMM, which features an easy-to-understand method for balancing the extensiveness and homogeneity of the clustering results. The experimental study has validated its effectiveness and claimed to have several advantages, such as GSDMM being quick to converge and able to handle the sparse and high-dimensional problem of short texts. GSDMM can also estimate the number of clusters automatically (Yin & Wang, 2014).

To address the limitation of DMM, that the single-topic assumption may be too strong for some datasets, Li et al. (2016) and Li et al. (2017) extended DMM with Poisson distribution to model the topic with a small number of topics, which is known as PDMM. Nonetheless, DMM and PDMM lack semantic relations between words, which were later improved by implementing the generalized Pólya urn (GPU) model. Through this, GPUDMM and GPUPDMM were developed to allow word semantic relations learned from millions of external documents (Li et al., 2017). Despite high computational costs, both GPUDMM and GPUPDMM methods have been proven to outperform DMM and individual PDMM methods (Murshed et al., 2022).

Nguyen et al. (2015) developed LFDMM by combining the Latent Feature vector and DMM, known to improve the feature word representations and incorporating word-topic mapping. The LF-DMM creates the words using either a latent feature model or a Dirichlet multinomial model, presuming that each text is sampled by a single topic. In other words, LFDMM extends Dirichlet multinomial models by combining latent feature vector representations of words learned on very large corpora to enhance the word-topic mapping trained on a smaller corpus (Nguyen et al., 2015). In addition, the DMM model can practically reach stability in 30 iterations and con-verge quickly to the best solutions (Qiang et al., 2019). Thus, this study aims to run experiments using these four topic modellings on anxiety-therapy text messages.

METHODOLOGY

The experiments were conducted on a Windows 10 Pro with Intel(R) Core (TM) i5-6200U CPU and 8GB RAM. The experiments were conducted to examine the best-performing topic modelling by comparing them in classifying text messages of anxiety therapy into different topics. An early topic modelling of Latent Dirichlet Allocation (LDA) introduced by Blei et al. (2003) has weaknesses in dealing with short text. As solutions, scholars introduced various improvised topic modelling techniques that are more applicable to short text (Nguyen et al., 2015; Yan et al., 2013). Four topic modellings, namely Latent Feature Dirichlet Multinomial Mixture (LFDMM), Gibbs Sampling Dirichlet Multinomial Mixture (GSDMM), Generalized Polya urn Dirichlet Multinomial Mixture (GPUDMM) and Poisson-based Dirichlet Multinomial Mixture Model (GPU-PDMM) were included in the initial experiment. For the purpose of examining the best-performing topic modelling, 28 text messages of anxiety therapy were used. Prior to this, a set of hyperparameters need to be set to execute the task which is explained in the next section.

Data Source

To compose a dataset for the current study, text messages used by mental health professionals in treating anxiety were collected from published studies between 2014 and 2021(Agyapong et al., 2016, 2017; Alfonsson et al., 2019; Anstiss & Davies, 2015; Chandran et al., 2019; Clough & Casey, 2018; Fitzpatrick et al., 2017; Furber et al., 2014; García et al., 2019). A dataset composed of 28 text messages of anxiety therapy were used. Text pre-processing is performed to prepare the text messages for analysis. The pre-processing steps includes removing stop words, numeral, punctuation, separator and special characters, converting all text to lowercase, and lemmatization. The statistics of the dataset include 248 words, 166 unique words, the maxi-mum length of a sentence is 16 words, 8.86 words per sentence on average, 1350 letters and 48.21 letters per sentence on average. Examples of text messages contained in the dataset include the following:

- “Patience is bitter but its result is sweeter. Have patience.”
- “Don’t be nervous, it’s all about finding ways that work for you.”
- “What lies behind you and what lies before you are tiny matters compared to what lies within you. Have faith in yourself and success can be yours.”

Parameter Settings

The current study runs the experiments using the open-source Short Text Topic Modelling (STTM) with a JAVA-based library (Qiang et al., 2019). Past literature was reviewed to select the hyperparameter values for all four short text topic modelling. Two important parameters that are used to control the topic inference are alpha and beta, where the alpha controls the pre-distribution of the weights of topics in each document, whereas the parameter of beta controls the pre-distribution of the weights of words in each topic. The higher the alpha value indicates that each document is likely to have a mixture of most topics (Gupta et al., 2022). In contrast to a low beta value, which indicates that the topic may be represented by few words, a high beta value indicates that each topic is likely to contain a mixture of more words, not just any word particularly (Gupta et al., 2022). Some studies used smaller α values ($\alpha = 0.05$) within the Latent Dirichlet Allocation (LDA) experiments (Schneider & Vlachos, 2018), while other studies used $\alpha = 0.1$ (Nguyen et al., 2015). Analysis using GSDMM and LFDMM also applied $\alpha = 0.1$ in other studies (Nguyen et al., 2015; Qiang et al., 2019). Other topic modelling, for instance DMM and GPUDMM, worked well when $\alpha = 50/K$ (Li et al., 2016; Qiang et al., 2019). Furthermore, when applied to short texts, several research studies also employed a smaller value of $\beta = 0.01$; (Nguyen et al., 2015; Qiang et al., 2019; Yan et al., 2013). However, other studies also used $\beta = 0.1$ for GSDMM (Qiang et al., 2019; Yin & Wang, 2014). Based on these pieces of literature, it is assumed that smaller values of hyperparameter settings are appropriate for short texts, where the smaller value of β is related to topic word-sparseness, while the smaller value of α implies that short text texts contain less topic (Yin & Wang, 2014). Since the data used in the current study are short text messages, thus, combinations of hyperparameters $\alpha = (0.05, 0.1)$ and $\beta = (0.01, 0.1)$ are experimented to identify the most suitable ones for the further analysis in the future.

All topic modelling methods were run for 2000 iterations. For LFDMM, where the iteration baseline model was run 1500 times. Then, the sample produced by the base-line model is used to initialize and run a further 500 iterations, as stated in the original paper (Nguyen et al., 2015). These word embedding-based topic modelling were trained using pre-trained word embeddings (“glove.6B.300d.txt”), where the dimension of the vector is 300. For LFDMM, λ is set to 0.6, which has shown in the past study that after combining data from a bigger external corpus with corpus-specific topic-word multinomials, better word-topic distributions were produced in this way.

Topic modelling assumes that there are K topics across all documents. In addition, the number K of topics needs to be determined to select the topic modelling that better fits data. Notably, a more significant number of topics may result in model over-fitting; thus, it is essential to be cautious while choosing the number of topics for training the model (Xie et al., 2021). To begin with, the current study hired a mental health professional to carefully and qualitatively examine all 28 text messages con-tent. Then, each text message was interpreted and labelled with a topic by the mental health professional based on her judgement and knowledge in the mental health therapy context. Overall, the results indicate eight labels for the anxiety-therapy text messages dataset. Hence, the number of topics K = 8 is used in the experiments. This results 4 models trained per method, as summarized in Table 1.

Table 1. Model Hyperparameter

Topic modelling	Hyperparameter
LFDMM	$\alpha = (0.05, 0.1)$, $\beta = (0.01, 0.1)$, $\lambda = 0.6$
GSDMM	$\alpha = (0.05, 0.1)$, $\beta = (0.01, 0.1)$
GPUDMM	$\alpha = (0.05, 0.1)$, $\beta = (0.01, 0.1)$, threshold = 0.5, weight = 0.1, filter size = 10
GPUPDMM	$\alpha = (0.05, 0.1)$, $\beta = (0.01, 0.1)$, $\lambda = 1.5$, threshold = 0.1, filter size = 10

Statistical Analysis

To select the topic modelling that better fits the data and assess the best performs topic modelling with combinations of hyperparameters and K = 8, classification ac-curacy, clustering and coherence score have been considered evaluation metrics. This performance measurements also were used in previous studies (Nguyen et al., 2015; Qiang et al., 2019). The dataset underwent five times of cross-validation to classify the accuracy. A higher accuracy demonstrates that the model is better at distinguishing the learned topics. Additionally, it suggests that the learned topics and the topics of the actual dataset are more closely aligned. In document clustering, every topic is treated as a cluster. Following the topic probability calculation, the topic (z) with the highest probability, p (z | d), was assigned to each document (d). The two most popular clustering measures, Purity and Normalized Mutual Information (NMI) are also applied in the experiment. A score close to 1 indicates better clustering, while a value close to 0 indicates weak clustering. Next, coherence evaluation is used to gauge the quality of topic-word distribution. A coherence score is based on the assumption that words with similar meanings would commonly coexist in a document. It should be highlighted that the coherence score accurately evaluates only frequently recurring terms in a document. This is because the less common word used to estimate the distribution of topic words is unreliable (Yan et al., 2013). In this experiment, a Wikipedia meta-document was used to determine the topic coherence score. The Point-Wise Mutual Information (PMI) of each word pair was calculated using a 10-word sliding window, accounting for word co-occurrence over the whole dataset of Wikipedia articles.

RESULT AND DISCUSSION

This section presents the result of testing four different topic modelling to assess whether it is possible to characterize text messages of anxiety therapy by topics. For each hyperparameter combination of $\alpha = (0.05, 0.1)$ and $\beta = (0.01, 0.1)$ with K = 8, four models were run 20 times for each topic modelling. The mean performance of the 20 runs were calculated to. Then, it reports the mean performance of classification accuracy, clustering, and topic coherence. Topic modelling algorithms are not deterministic. This means that it could produce different results when the same algorithm is run on the same dataset due to its behavior incorporating elements of randomness and probability. In turn, the model may make slightly different predictions. It may have a slightly different performance when evaluated. The mean performance of classification accuracy, clustering, and topic coherence are related to the mean achieved across twenty different runs. These twenty different runs using the same dataset and the experiment test how consistently the algorithm performs. Therefore, the mean is used as a measurement in this experiment

Classification Accuracy

The mean classification accuracy score of all four topic modelling methods with four different hyperparameter settings and 8 number of topics were illustrated in Figure 1. Most of the methods achieved accuracy scores close to each other when K = 8 with all combinations of hyperparameters except for GPUPDMM when α = 0.1 and β = 0.01. In that situation, it is noticeable that GPUPDMM shows a remarkably lowest score across the experiments compared to the rest of the methods. Nonetheless, GPUPDMM has almost the identical accuracy scores when applied to combinations of α = (0.05, 0.1) and β = (0.1). Among all hyperparameter combinations, LFDMM scores the highest accuracy twice compared to the rest of the methods. In addition, the highest accuracy score was achieved by LFDMM when α = 0.1 and β = 0.01. GSDMM shows the same up-and-down pattern as LFDMM and performs the best classification accuracy with smaller values of alpha and beta. Of all four combinations of hyperparameters, GPUPDMM and GPUDMM achieve the highest score when a bigger value of alpha and beta is applied (α = 0.1 and β = 0.1). As comparisons, GPUDMM and GSDMM have almost consistent accuracy performance across the four different combinations of hyperparameters. Furthermore, all methods show comparatively small accuracy differences between each other when α = 0.05 and β = 0.1.

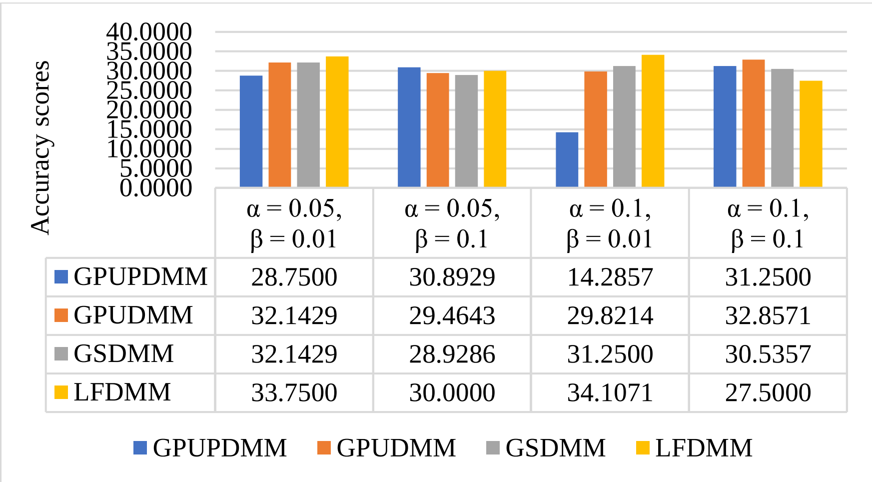

Figure 1: Mean classification of accuracy of four topic modelling

Clustering

The data in Table 2 report the average of the Purity and Normalized Mutual In-formation (NMI) score of each modelling on the anxiety dataset. The closer the value of Purity and NMI scores to 1 indicates better clustering performance, whereas the closer the value to 0 indicates poorer clustering performance (Nguyen et al., 2015). The highest score of Purity and NMI of each modelling which comprises of different hyperparameter setting are highlighted in bold. It can be observed that GPUDMM achieves the lowest Purity and NMI scores amongst all models for anxiety dataset. On the other hand, LFDMM scores the highest in clustering, followed by GSDMM and GPUPDMM. It is worth noting that both GSDMM and LFDMM performed the best clustering when α = 0.1 and β = 0.01. Other than that, three methods scored the highest in purity evaluation when α = 0.1, whereas all methods performed the best in NMI evaluation when β = 0.01.

Table 2. Purity and NMI scores of four topic modelling

		GPUPDMM		GPUDMM		GSDMM		LFDMM	
alpha	**beta**	**Purity**	**NMI**	**Purity**	**NMI**	**Purity**	**NMI**	**Purity**	**NMI**
0.05	<u>0.01</u>	0.4429	**0.4332**	**0.3464**	**0.2231**	0.4536	0.4426	0.4893	0.4787
0.05	0.1	0.4443	0.4270	0.3304	0.1834	0.4500	0.4396	0.4643	0.4436
<u>0.1</u>	<u>0.01</u>	0.4286	0.4031	0.3143	0.1809	**0.4607**	**0.4448**	**0.5000**	**0.4808**
<u>0.1</u>	0.1	**0.4446**	0.4329	0.3071	0.1619	0.4518	0.4447	0.4625	0.4409

Coherence

The overview of coherence scores of each topic modelling method is shown in Figure 2. Overall, GPUDMM deviates to the lowest scores from the rest of the methods for all combinations of the hyperparameters. Moreover, all methods deviate from each other mostly when $\alpha = 0.05$ and $\beta = 0.1$. This indicates that these hyperparameters are less reliable for drawing relevant topics. Other than GPUDMM, the other three methods' coherence scores are close to each other when $\alpha = 0.05$ and $\beta = 0.01$, $\alpha = 0.1$ and $\beta = 0.01$, as well as $\alpha = 0.1$ and $\beta = 0.1$. In other words, these topic modelling can infer more relevant topics when these parameter values are in use. It is also worth to noting that as in accuracy evaluation, LFDMM again scores the highest twice in coherence evaluation with the same set of parameters setting.

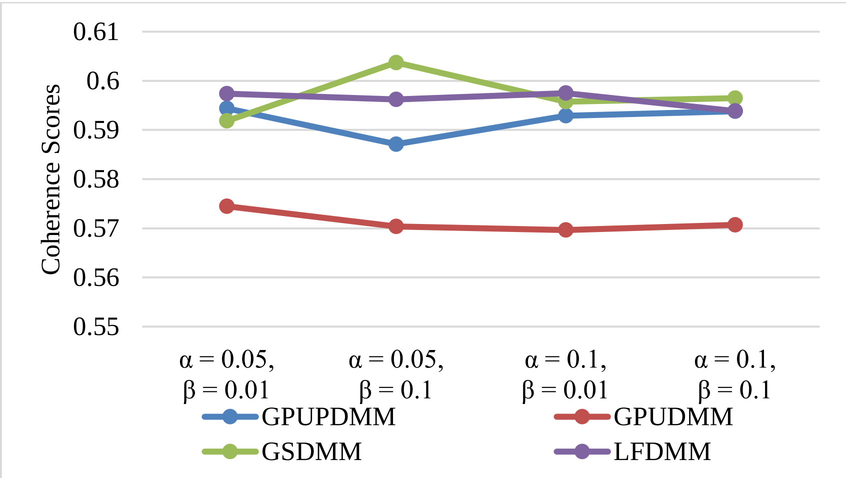

Figure 2: Coherence scores of four topic modelling

Each topic modelling method performs differently with different combinations of alpha and beta values. Due to the accuracy scores of most methods being close to each other and the differences between values being comparatively small, it is difficult to recognize which method should be considered the most suitable method for the anxiety-therapy text messages dataset when K = 8. The comparatively small difference achieved by all methods across all hyperparameters, values indicates that all methods' performances depend slightly on the hyperparameter setting except for GPUPDMM. A steep decrease in accuracy evaluation shown by GPUPDMM when $\alpha = 0.1$ and $\beta = 0.01$ indicates that the hyperparameters influence this method's performance.

The aim of comparing the distinct topic modelling methods is to find out which combination of parameter settings best favor the performance accuracy for anxiety-therapy text messages. Although the current experiment shows that the accuracy performances of applied topic modelling are not so strong in making classification correctly, the absolute values of the metrics indicate that the methodologies applied are still able to find meaningful hidden topics and worthy of further exploration with a bigger sample dataset. The result emphasizes the importance of designing topic modelling experiments that support therapist and patients not only in content com-prehension but also in research, treatment planning, and therapy text message analysis. Overall, the evaluation of topic accuracy shows that LFDMM, on average, pro-duces more accurate topics and this method is hyperparameter setting dependent just like the rest of the topic modelling methods. For some values of alpha and beta, GPUPDMM and GPUDMM achieved higher accuracy scores than other topic model-ling methods, but LFDMM outperformed the rest of the models twice. As mentioned before, LFDMM is hyperparameter setting dependent. Thus, it is good to know further which combination of alpha and beta work best together with LFDMM to run the anxiety-therapy text messages using a bigger sample dataset in the future.

The lowest score achieved by GPUDMM in clustering indicates that GPUDMM could not label the sample of text messages into its cluster correctly when compared to the golden label. In contrast, LFDMM did not just outperform the rest of the model in accuracy but also in clustering evaluations. Both the highest accuracy and clustering scores were achieved by LFDMM when $\alpha = 0.1$ and $\beta = 0.01$ were applied. Furthermore, a higher value of alpha (0.1) was found to contribute to higher scores in purity evaluation for three methods, which indicates that the anxiety-therapy text messages are likely to have more than one topic. It is possible due to the fact that some people with anxiety also experience other related mental health problems such as depression and alcohol misuse disorder (Agyapong et al., 2020). Meanwhile, the highest NMI scores in clustering evaluations shown by all methods proved that low beta value worked well with short text data as the topic in the dataset of anxiety-therapy was represented by few words. Although the resulting value of NMI scores interpreted the data as not perfectly clustered, the NMI score for LFDMM, which is close to 0.5, indicates that almost 88% of the data are correctly clustered (Kachouie & Shutaywi, 2020). This finding could further underline the exploration with more extensive dataset of text messages to estimate related anxiety-therapy topics and words. It can be beneficial for mental health professionals practicing online therapy for mental health disorders to be more proactive towards patients' treatment plans and enhance health communication skills.

Other than GPUDMM, the other three methods' coherence scores are close to each other when $\alpha = 0.05$ and $\beta = 0.01$, $\alpha = 0.1$ and $\beta = 0.01$, as well as $\alpha = 0.1$ and $\beta = 0.1$. The underlined alpha and beta show that the same values applied in those models have contributed to finding relevant topics. Based on these explanations, the combi-nations of hyperparameters $\alpha = 0.1$ and $\beta = 0.01$ are the most suitable for working with LFDMM and drawing topics for anxiety-therapy text messages.

By using the Dirichlet-Multinomial distribution as the foundational model, the Dirichlet-Multinomial Mixture (DMM) improves the ability to capture dependencies between words within individual topics compared to LDA. This leads to more accurate and realistic topic modeling, particularly in scenarios involving sparse or noisy data. The DMM is a more flexible approach, well-suited for tasks where documents are composed of a mixture of multiple topics, such as in the analysis of anxiety therapy content.

Many existing models, such as LDA and k-means clustering, primarily emphasize optimizing an objective function, yet they often do not prioritize the accuracy with which the resulting topics or clusters represent the true underlying structure of the data. In contrast, the use of PMI and coherence evaluations in this study enhances the human interpretability of the model. Coherence ensures that topics are not arbitrary groupings of words but are instead semantically meaningful. From a practical relevance standpoint, improved coherence facilitates the generation of more action-able insights. For instance, in the context of topic modeling applied to anxiety therapy text messages, a coherent set of topics uncovered distinct and meaningful theme.

CONCLUSION

In conclusion, it is challenging to determine the superiority of one method over an-other because most methods yield similar results. When selecting a topic modeling approach for analyzing anxiety-therapy text messages, one should consider both its effectiveness and efficiency. In the present study, LFDMM demonstrates the highest accuracy scores at $\alpha = 0.1$ and $\beta = 0.01$, but it requires more time and is more expensive compared to other methods. All methods show nearly identical accuracy at $\alpha = 0.05$ and $\beta = 0.1$, suggesting that these models have similar capabilities in generating topic-document distribution with the specified hyperparameter combinations. This makes it difficult to choose the most suitable topic modeling approach for future experiments on anxiety-therapy text messages. Therefore, it is strongly recommended to experiment with different hyperparameter combinations, as well as to incorporate clustering and coherence evaluations in the initial experiment, in order to determine the most appropriate setting for future use.

REFERENCES

Aguilera, A., Hernandez-Ramos, R., Haro-Ramos, A. Y., Boone, C. E., Luo, T. C., Jing, X., Chakraborty, B., Karr, C., Darrow, S., & Figueroa, C. A. (2021). A Text Messaging Intervention (StayWell at Home) to Counteract Depression and Anxiety During COVID-19 Social Distancing : Pre-Post Study. JMIR Mental Health, 8(11). https://doi.org/10.2196/25298

Agyapong, V. I. O., Hrabok, M., Vuong, W., Gusnowski, A., Shalaby, R., Surood, S., Greenshaw, A. J., Aulakh, A., Abba-Aji, A., & Singh, M. (2020). Implementation and evaluation of a text message-based addiction counseling program (text4hope-addiction support): Protocol for a questionnaire study. JMIR Research Protocols, 9(11). https://doi.org/10.2196/22047

Agyapong, V. I. O., Juhás, M., Ohinmaa, A., Omeje, J., Mrklas, K., Suen, V. Y. M., Dursun, S. M., & Greenshaw, A. J. (2017). Randomized controlled pilot trial of supportive text messages for patients with depression. BMC Psychiatry, 17(1), 1–10. https://doi.org/10.1186/s12888-017-1448-2

Agyapong, V. I. O., Mrklas, K., Juhás, M., Omeje, J., Ohinmaa, A., Dursun, S. M., & Greenshaw, A. J. (2016). Cross-sectional survey evaluating Text4Mood: Mobile health program to reduce psychological treatment gap in mental healthcare in Alberta through daily supportive text messages. BMC Psychiatry, 16(1), 1–12. https://doi.org/10.1186/s12888-016-1104-2

Alfonsson, S., Englund, J., & Parling, T. (2019). Tailored text message prompts to increase therapy homework adherence: A single-case randomised controlled study. Behaviour Change, 36(3), 180–191. https://doi.org/10.1017/bec.2019.10

Anstiss, D., & Davies, A. (2015). “Reach Out, Rise Up”: The efficacy of text messaging in an intervention package for anxiety and depression severity in young people. Children and Youth Services Review, 58, 99–103. https://doi.org/10.1016/j.childyouth.2015.09.011

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. Journal of Machine Learning Research, 3(4–5), 993–1022. https://doi.org/10.1016/b978-0-12-411519-4.00006-9

Chandran, S., Mathur, S., & Rao, K. M. (2019). Media and the role of digital psychiatry in mental health. Digital Psychiatry, 2(1), 34–44. https://doi.org/10.1080/2575517x.2019.1651625

Clough, B. A., & Casey, L. M. (2018). Will patients accept daily SMS as a communication to support adherence to mental health treatment? Daily SMS: Acceptance, feasibility, & satisfaction. International Journal of Cyber Behavior, Psychology and Learning, 8(3), 24–35. https://doi.org/10.4018/IJCBPL.2018070103

Fairie, P., Zhang, Z., D’Souza, A. G., Walsh, T., Quan, H., & Santana, M. J. (2021). Categorising patient concerns using natural language processing techniques. BMJ Health and Care Informatics, 28(1), 1–9. https://doi.org/10.1136/bmjhci-2020-100274

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. JMIR Mental Health, 4(2), 1–11. https://doi.org/10.2196/mental.7785

Furber, G., Jones, G. M., Healey, D., & Bidargaddi, N. (2014). A comparison between phone-based psychotherapy with and without text messaging support in between sessions for crisis patients. Journal of Medical Internet Research, 16(10), e219. https://doi.org/10.2196/jmir.3096

García, Y., Ferrás, C., Rocha, Á., & Aguilera, A. (2019). Exploratory Study of Psychosocial Therapies with Text Messages to Mobile Phones in Groups of Vulnerable Immigrant Women. Journal of Medical Systems, 43(8). https://doi.org/10.1007/s10916-019-1393-3

Gupta, R. K., Agarwalla, R., Naik, B. H., Evuri, J. R., Thapa, A., & Singh, T. D. (2022). Prediction of research trends using LDA based topic modeling. Global Transitions Proceedings, 3(1), 298–304. https://doi.org/10.1016/j.gltp.2022.03.015

Kachouie, N. N., & Shutaywi, M. (2020). Weighted mutual information for aggregated kernel clustering. Entropy, 22(3), 1–15. https://doi.org/10.3390/e22030351

Kao, S. W., & Luarn, P. (2020). Topic modeling analysis of social enterprises: Twitter evidence. Sustainability (Switzerland), 12(8), 1–20. https://doi.org/10.3390/SU12083419

Li, C., Duan, Y., Wang, H., Zhang, Z., Sun, A., & Ma, Z. (2017). Enhancing topic modeling for short texts with auxiliary word embeddings. ACM Transactions on Information Systems, 36(2), 1–30. https://doi.org/10.1145/3091108

Li, C., Wang, H., Zhang, Z., Sun, A., & Ma, Z. (2016). Topic modeling for short texts with auxiliary word embeddings. SIGIR 2016 - Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, 165–174. https://doi.org/10.1145/2911451.2911499

Murshed, B. A. H., Mallappa, S., Abawajy, J., Saif, M. A. N., Al-ariki, H. D. E., & Abdulwahab, H. M. (2022). Short text topic modelling approaches in the context of big data: taxonomy, survey, and analysis. Artificial Intelligence Review. https://doi.org/10.1007/s10462-022-10254-w

Nguyen, D. Q., Billingsley, R., Du, L., & Johnson, M. (2015). Improving Topic Models with Latent Feature Word Representations. Transactions of the Association for Computational Linguistics, 3, 598–599. https://doi.org/10.1162/tacl_a_00245

Qiang, J., Qian, Z., Li, Y., Yuan, Y., & Wu, X. (2019). Short text topic modeling techniques, applications, and performance: A survey. Journal of Latex Class Files, 14(8), 1–17. https://doi.org/10.1109/tkde.2020.2992485

Ritvo, P., Knyahnytska, Y., Pirbaglou, M., Wang, W., Tomlinson, G., Zhao, H., Linklater, R., Bai, S., Kirk, M., Katz, J., Harber, L., & Daskalakis, Z. (2021). Online mindfulness-based cognitive behavioral therapy intervention for youth with major depressive disorders: Randomized controlled trial. Journal of Medical Internet Research, 23(3), 1–11. https://doi.org/10.2196/24380

Schneider, J., & Vlachos, M. (2018). Topic modeling based on keywords and context. Proceedings of the 2018 SIAM International Conference on Data Mining, May, 369–377. https://doi.org/10.1137/1.9781611975321.42

Voskergian, D., Bakir-Gungor, B., & Yousef, M. (2023). TextNetTopics Pro, a topic model-based text classification for short text by integration of semantic and document-topic distribution information. Frontiers in Genetics, 14(October), 1–23. https://doi.org/10.3389/fgene.2023.1243874

Weisser, C., Gerloff, C., Thielmann, A., Python, A., Reuter, A., Kneib, T., & Säfken, B. (2023). Pseudo-document simulation for comparing LDA, GSDMM and GPM topic models on short and sparse text using Twitter data. Computational Statistics, 38(2), 647–674. https://doi.org/10.1007/s00180-022-01246-z

Xie, R., Chu, S. K. W., Chiu, D. K. W., & Wang, Y. (2021). Exploring Public Response to COVID-19 on Weibo with LDA Topic Modeling and Sentiment Analysis. Data and Information Management, 5(1), 86–99. https://doi.org/10.2478/dim-2020-0023

Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013). A biterm topic model for short texts. WWW 2013 - Proceedings of the 22nd International Conference on World Wide Web, May, 1445–1455. https://doi.org/10.1145/2488388.2488514

Yin, J., & Wang, J. (2014). A Dirichlet multinomial mixture model-based approach for short text clustering. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 233–242. https://doi.org/10.1145/2623330.2623715

Zuo, Y., Wu, J., Xhang, H., Lin, H., Wang, F., Xu, K., & Xiong, H. (2016). Topic Modeling of Short Texts: A Pseudo-Document View. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2105–2114. https://doi.org/10.1109/TKDE.2021.3073195

Zuo, Y., Zhao, J., & Xu, K. (2016). Word network topic model: a simple but general solution for short and imbalanced texts. Knowledge and Information Systems, 48(2), 379–398. https://doi.org/10.1007/s10115-015-0882-z