






Article

BERTopic for Enhanced Idea Management and Topic Generation in Brainstorming Sessions

Asma Cheddak ^{1,†} , Tarek Ait Baha ^{2,†} , Youssef Es-Saady ^{2,3,*} , Mohamed El Hajji ^{2,4,*} 
and Mohamed Baslam ¹ 

- ¹ TIAD Laboratory, Sultan Moulay Slimane University, Beni-Mellal 23000, Morocco; asmae.cheddak@usms.ma (A.C.); m.baslam@usms.ma (M.B.)
² IRF-SIC Laboratory, Ibnou Zohr University, Agadir 80000, Morocco; t.aitbaha@uiz.ac.ma
³ Polydisciplinary Faculty of Taroudant, Ibnou Zohr University, Taroudant 83000, Morocco
⁴ Regional Center for Education and Training Professions-Souss Massa, Agadir 80000, Morocco
* Correspondence: y.essaady@uiz.ac.ma (Y.E.-S.); m.elhajji@crmefsm.ac.ma (M.E.H.)
† These authors contributed equally to this work.

Abstract: Brainstorming is an important part of the design thinking process since it encourages creativity and innovation through bringing together diverse viewpoints. However, traditional brainstorming practices face challenges such as the management of large volumes of ideas. To address this issue, this paper introduces a decision support system that employs the BERTopic model to automate the brainstorming process, which enhances the categorization of ideas and the generation of coherent topics from textual data. The dataset for our study was assembled from a brainstorming session on “scholar dropouts”, where ideas were captured on Post-it notes, digitized through an optical character recognition (OCR) model, and enhanced using data augmentation with a language model, GPT-3.5, to ensure robustness. To assess the performance of our system, we employed both quantitative and qualitative analyses. Quantitative evaluations were conducted independently across various parameters, while qualitative assessments focused on the relevance and alignment of keywords with human-classified topics during brainstorming sessions. Our findings demonstrate that BERTopic outperforms traditional LDA models in generating semantically coherent topics. These results demonstrate the usefulness of our system in managing the complex nature of Arabic language data and improving the efficiency of brainstorming sessions.

Keywords: brainstorming; support system; BERTopic; clustering; topic modeling



Citation: Cheddak, A.; Ait Baha, T.; Es-Saady, Y.; El Hajji, M.; Baslam, M. BERTopic for Enhanced Idea Management and Topic Generation in Brainstorming Sessions. *Information* **2024**, *15*, 365. <https://doi.org/10.3390/info15060365>

Academic Editor: María N. Moreno García

Received: 12 May 2024
Revised: 8 June 2024
Accepted: 17 June 2024
Published: 20 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Brainstorming is an essential part of the design thinking process and plays a key role in generating ideas and solving problems, especially in large organizations. Traditional brainstorming stimulates creativity and innovation by bringing together varied viewpoints and skills, which is critical for tackling complex corporate situations. However, despite their inherent value, traditional methods encounter significant challenges when applied in contemporary organizational settings.

In today’s changing business setting, which is characterized by globalization and remote work arrangements, traditional brainstorming encounters numerous obstacles. These include managing multiple participants from different locations, overcoming language barriers, and dealing with time constraints. The manual process of idea sorting and analysis further complicates these challenges, often resulting in inefficiencies and subjectivity in idea selection. Moreover, the increasing prevalence of remote work necessitates the adoption of digital platforms for collaborative brainstorming sessions. While existing technologies facilitate remote idea generation [1,2], many lack the capability to automatically categorize and analyze brainstorming outputs effectively [3]. This limitation restricts the extraction

of useful insights from the amount of data collected during these sessions, limiting informed decision-making and innovation within organizations. A brainstorming session goes through several phases: idea generation, the grouping and structuring of ideas, idea evaluation, deliberation, and sharing of the most important results. Many studies have demonstrated the effectiveness of using artificial intelligence in the initial phase of idea generation [1,2]. Our contribution mainly focuses on the second phase. It aims to assist the facilitator in categorizing and structuring ideas more effectively, free from subjectivity or group influence. While the facilitator is indeed crucial in brainstorming, the need for digital brainstorming has become evident to help the facilitator manage current challenges. Welcoming ideas from a large group, structuring them, and presenting them effectively in minimal time is a daunting task, often impossible for a human. Hence, technology intervention is crucial to optimizing brainstorming.

Recognizing these challenges, our research aims to address the shortcomings of traditional brainstorming methods by proposing an automated system that accelerates the brainstorming process and increases its usefulness in management and business contexts. In our preliminary work [4], we presented a system designed to recognize ideas written by participants on sticky notes during brainstorming sessions. This system was designed to identify domain-specific entities embedded within these notes. Our earlier efforts specifically focused on indexing these extracted ideas into a knowledge base, ensuring their accessibility and relevance for future reference and application.

Based on this foundation, the current work represents a significant advancement and extension of our previous efforts using BERTopic. While our initial system focused on capturing and organizing brainstorming ideas, this study aims to address issues related to the analysis and categorization of these ideas. The BERTopic model is integrated into our system to automatically generate topics from the collected ideas, thus facilitating the understanding of the brainstorming outputs. Preliminary evaluation results indicate that the proposed method is effective in generating coherent and meaningful topics from brainstorming session data.

As a key contribution of our paper, we introduce the novel application of BERTopic to automate the process of brainstorming. To the best of our knowledge, this is the first instance of using BERTopic for this purpose. Through this approach, we present a novel method to improve the efficacy of brainstorming sessions. Additionally, we offer a transformative solution to modern businesses' changing demands by bridging the gap between traditional brainstorming methods and contemporary technological advances.

The rest of the paper is organized as follows: Section 2 covers background knowledge and related work. It covers brainstorming systems and also explores existing topic modeling techniques and related work. Section 3 presents our proposed method. Section 4 focuses on experimentation and findings, detailing the dataset and showing the experimental results. Finally, in Section 5, we conclude the paper with a summary of our contributions and our future research.

2. Related Works

The practice of brainstorming has been in use for as long as humans have lived in communities. Throughout history, people have gathered to find solutions to problems threatening their well-being as part of collective resolution. The term "brainstorming" was first used in around 1948 by Osborn, marking the beginning of this innovative practice to solve the majority of economic problems, particularly in management or marketing [5].

Many methods have been developed over time, but no study to date has shown the effectiveness of one over the other. Generally, brainstorming can be classified into the following types:

- Traditional group brainstorming: This is the classic process where a group of people physically gather to generate ideas interactively and collaboratively.

- Virtual group brainstorming: This method uses digital tools such as online brainstorming software to allow participants to contribute remotely in real time or asynchronously.
- Individual electronic brainstorming: Participants generate ideas individually using electronic tools, then share them with the group for discussion and evaluation.
- Hybrid brainstorming: This approach combines elements of traditional group brainstorming with electronic tools to facilitate the collection, categorization, and presentation of ideas.

Traditional brainstorming sessions have served as a cornerstone for collaborative problem solving and innovation. These sessions consist of group discussions where participants are encouraged to express their ideas without fear of judgment. Despite its extensive use, traditional brainstorming has encountered certain challenges. One notable limitation is the volume of ideas generated during these sessions [3]. As participants contribute their thoughts, the sheer number of ideas can become overwhelming, making it difficult for facilitators and team members to effectively manage and analyze these outputs. Additionally, the manual process of sorting these ideas can be time-consuming and subjective.

Studies have been conducted on comparing the effectiveness of virtual and traditional brainstorming [6]. They show that virtual brainstorming offers advantages such as ease of idea collection and categorization, as well as remote participation. These studies generally encourage combining traditional and digital methods to maximize efficiency and promote creativity among participants [5]. The manual process of sorting through, categorizing, and analyzing the ideas produced during these sessions is both time-consuming and subjective. Moderators, facilitators, and team leaders must filter through a large number of ideas and proposals, which requires significant time and effort. The subjective nature of this process means that the selection and prioritization of ideas are influenced by the facilitator's personal biases and perceptions, which may not always align with the collective's objectives or the outcomes intended from the session. Therefore, the increasing popularity of remote work has made digital platforms essential for brainstorming sessions [7]. This change affects the way ideas are shared, collected, and managed. While several technologies and platforms have been developed to facilitate remote brainstorming and idea management [8], most existing solutions lack the capability to automatically categorize and analyze the brainstorming outputs, potentially missing opportunities to extract meaningful insights from the data collected [3]. Recognizing these limitations, there has been increasing interest in using technology to supplement and simplify the brainstorming process. Advancements in digital collaboration tools, such as online brainstorming platforms and virtual whiteboards, have facilitated the transition from traditional sessions to remote settings [9,10]. These technologies offer features such as real-time idea sharing and asynchronous collaboration, which enhance the flexibility of brainstorming sessions. Recently, the integration of machine learning and natural language processing (NLP) techniques has created new opportunities for automating various aspects of the brainstorming process [11]. These techniques are used to organize and analyze textual documents in various application domains [12].

Topic modeling has evolved significantly since its inception, with early methods like Latent Semantic Analysis (LSA) [13] and Probabilistic Latent Semantic Analysis (pLSA) [14] paving the way for more advanced techniques [15]. Latent Dirichlet Allocation (LDA) [16,17] became a standard, introducing a generative probabilistic model that assumes documents are mixtures of topics, which in turn are mixtures of words. However, LDA's reliance on Dirichlet priors and the need for pre-specification of the number of topics have led to the development of non-parametric methods and the incorporation of deep learning to improve topic modeling. When compared to traditional methods such as LDA and NMF, BERTopic has been found to perform exceptionally well, particularly in datasets with a large number of short texts. Studies have shown that BERTopic can uncover related topics around specific terms, providing deeper insights [18]. It also allows for hierarchical topic reduction and has advanced search and visualization capabilities [19]. However,

for longer texts or smaller datasets, LDA and NMF may still be preferable due to their simplicity and lower resource requirements.

BERTopic is an algorithm that leverages BERT embeddings and clustering to perform topic modeling. When integrated into brainstorming support systems, it offers facilitators the ability to automatically model topics from the textual content generated during brainstorming sessions. This integration allows for the real-time analysis of discussion themes, helping participants to identify patterns and focus on areas with the most potential for innovation. This integration offers several advantages, including the use of contextual embeddings, which enhance semantic understanding, allowing for a deeper comprehension of the text and more precise topic representations. Additionally, the topic modeling feature enables continued updates of topic models as new ideas emerge during brainstorming sessions. The use of algorithms such as Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [20,21], which is well known for its ability to cluster data without needing the number of clusters as input, improves clustering within topic modeling with BERTopic, further enhancing the exploration and comprehension of brainstorming concepts. Furthermore, BERTopic provides a range of visualization tools, including topic hierarchies and inter-topic distance maps, which facilitate clearer insights and decision-making processes.

Table 1 provides a comprehensive overview of various topic modeling methods, including their advantages, limitations, interpretability, and computational complexity. The methods included are LDA, NMF, LSA, PLSA, and BERTopic. LDA is highly interpretable and capable of handling large datasets, though it may lack semantic representation. NMF offers interpretable results and effective dimensionality reduction, but it is sensitive to noise and requires preprocessed data. LSA employs semantic structures and is effective with large corpora, but it is less interpretable and sensitive to noise. PLSA captures nonlinear relationships through the use of a sophisticated probabilistic model; however, its implementation is complex and prone to overfitting. BERTopic employs BERT for contextual representation, capturing complex semantic relationships and providing precise results for large datasets, though it is resource-intensive and less interpretable than other techniques. This detailed comparison enables the identification of the trade-offs involved in selecting a suitable topic modeling technique, taking into account specific requirements such as dataset size, interpretability, and computational resources. Justifying the use of advanced methods such as BERTopic can strengthen the evaluation by emphasizing their effectiveness in contexts demanding nuanced semantic understanding and complex relationship modeling, despite their higher computational demands.

Table 1. Overview of topic modeling methods

Method	Advantages	Limitations	Interpretability	Computational Complexity
LDA [16]	Easy to interpret Handles large datasets	Simplified model May lack semantic representation	High	Moderate
NMF [22]	Interpretable results Effective for dimensionality reduction	Sensitive to noise Requires preprocessed data	Moderate	Moderate
LSA [13]	Uses semantic structure Works well for large corpora	Sensitivity to noisy data Less interpretable than LDA	Moderate	Moderate

Table 1. Cont.

Method	Advantages	Limitations	Interpretability	Computational Complexity
PLSA [14]	Sophisticated probabilistic model Captures nonlinear relationships	More complex to implement Sensitivity to overfitting	Moderate	High
BERTopic [23]	Utilizes BERT for contextual representation Captures complex semantic relationships Precise results for large datasets	Resource-intensive Less interpretable than traditional approaches Requires learning phase on large data	Moderate	High

3. Materials and Methods

This section provides a concise description of the proposed methodological approach, which is structured into several sequential layers. The method, centered around the BERTopic model, is designed to automatically generate topics from collected ideas. As shown in Figure 1, which presents the workflow of our approach, the process involves three primary steps: data preparation, topic modeling, and evaluation.

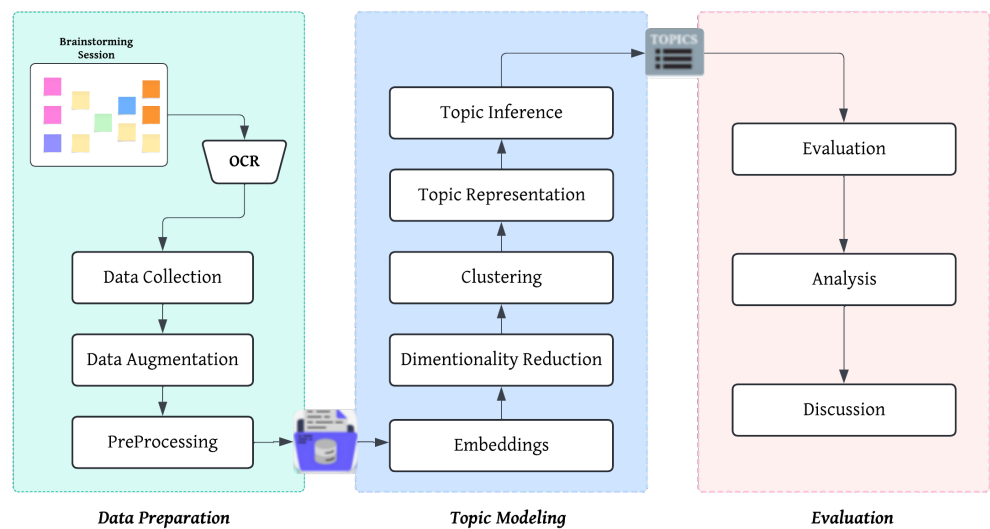


Figure 1. An overview of the methodological approach.

3.1. Data Preparation

In this section, we outline the process used to construct and enrich the database for our study.

3.1.1. Database Construction

The database for our study on thematic brainstorming around the issue of “scholar dropouts” was constructed using a methodical process consisting of several phases. Initially, data were collected from an organized brainstorming session on the topic of academic failure. During the session, participants were prompted to generate ideas and thoughts on the topic and record them on provided Post-it notes. At the end of the session, the Post-its were collected, which constituted the corpus of data for the study. Figure 2 presents an examples of ideas generated during a brainstorming session. The left part of the figure

shows ideas generated written in Arabic, while the right part provides their corresponding English translations.



Figure 2. Examples of ideas generated during the brainstorming session and their corresponding translations. (Left) Ideas generated during a brainstorming session written in Arabic. (Right) The corresponding English translations of ideas.

Afterwards, the handwritten data on the Post-it notes were digitized through an optical character recognition (OCR) process, which enabled the content to be converted into digital text. This step facilitated the subsequent automated processing of the data. To guarantee the quality of the extracted text, a postprocessing step was implemented. This involved applying a spelling checker to detect and correct any spelling or typographical errors.

The resulting ideas were then saved in a CSV format, ready for subsequent analysis. The given data were utilized to apply advanced natural language processing (NLP) and topic modeling techniques, which allowed for a thorough exploration of the ideas generated during brainstorming sessions. The methodology used to construct the database facilitated the conversion of raw brainstorming data into digital data that could be analyzed as part of this study. Figure 3 shows the process of database creation.

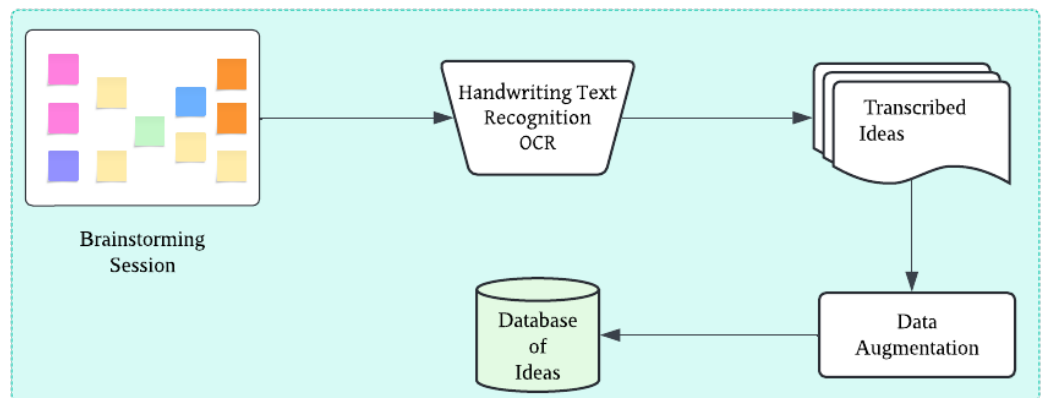


Figure 3. Process of database creation.

3.1.2. Data Augmentation

To expand our database and improve our initial dataset, we used a data augmentation approach based on the GPT-3.5 language model [24]. This method was used to introduce greater variability into our data and enhance the robustness of our future prediction models. Using the GPT model, we generated new idea samples by conditioning the model on our existing dataset. This approach allowed us to create synthetic variations of the original ideas while maintaining the semantic coherence with the theme of “scholar dropouts”.

In fact, we started with a database generated during a real brainstorming session. During this session, participants proposed 60 ideas, which were classified according to the estimations of the facilitator and the group. Based on this categorization, we utilized GPT-3.5 to generate comparable concepts but with various formulations. This is how the final database had 1080 samples.

Once we generated the augmented data, we integrated it into our existing database, resulting in a significant increase in both size and diversity. This augmentation process sets the stage for more robust modeling and enables deeper exploration of the relationships between different concepts related to “scholar dropouts”.

3.1.3. Data Preprocessing

Before starting our experiments, we performed a series of preprocessing steps to prepare the data. This preparation was essential for optimizing the analysis process and enhancing the accuracy of our findings. The preprocessing included a comprehensive review and cleanup of the dataset. Several Python libraries are available for NLP preprocessing. We chose NLTK (<https://www.nltk.org/>, accessed on 15 April 2024) for its simplicity and extensive resources. The data preprocessing started with the normalization and cleaning of the data. The second phase was the elimination of stopwords. To ensure the elimination of irrelevant terms, we used NLTK’s stopwords list for Arabic as well as an additional custom set.

3.2. Topic Modeling and Clustering

For our investigation, we used BERTopic due to its cutting-edge advancements and efficacy in topic modeling, as detailed in recent research [18,23]. BERTopic distinguishes itself through its modularity, which allows for a personalized approach to topic modeling and analysis. The process proceeds in several steps, beginning with dimensionality reduction, followed by the clustering of data into distinct topics using a specified algorithm (Algorithm 1), then proceeding to the analysis and visualization of outcomes, resulting in the ability to make predictions on novel data.

The BERTopic process begins by transforming our input texts into numerical representations, a crucial step for which several methods are available. However, we use sentence-transformers [25], chosen for their efficiency in generating embeddings. These models perform well at optimizing semantic similarity, which significantly improves the clustering task. Given our focus on Arabic data, we opted for “AraBERT-v02”, a transformer model provided by Hugging Face (<https://huggingface.co/models>, accessed on 15 April 2024). This choice enabled us to capture the semantic nuances of our text data, thus enhancing the accuracy and relevance of the generated clusters.

3.2.1. Dimensionality Reduction

A key feature of BERTopic lies in its ability to tackle the high dimensionality of input embeddings through dimensionality reduction. Embeddings, which are known for their dense representation, suffer from the curse of dimensionality, which complicates the clustering process. BERTopic addresses this challenge by compressing the embeddings into a more manageable space, which improves the efficiency of clustering methods. For our solution, we used Uniform Manifold Approximation and Projection (UMAP) to achieve this important step, taking advantage of its ability to preserve the local structure of data while reducing dimensionality [26].

Algorithm 1: BERTopic Process for Topic Modeling and Clustering [27]**Input:**

- A set of textual data
- Pre-trained transformer model for embedding generation (e.g., “AraBERT-v02”)

Output:

- Identified topics with key terms and visualizations

Steps:

1. Embedding Generation
 - Convert input texts into numerical representations (embeddings).
 - Model used: AraBERT-v02.
2. Dimensionality Reduction
 - Apply Uniform Manifold Approximation and Projection (UMAP) to reduce the dimensionality of embeddings.
3. Clustering
 - Use Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) to cluster the reduced embeddings.
4. Topic Representation
 - Extract topics from the identified clusters using c-TF-IDF.
5. Visualization
 - Generate visual representations of topics and key terms.

3.2.2. Clustering

After reducing the dimensionality of our input embeddings, the next step is to cluster them into groups of similar embeddings to extract our topics. This process of clustering is crucial because the effectiveness of our topic representations depends on the performance of our clustering technique. Therefore, employing a high-performing clustering technique ensures more accurate topic extraction from the data. Specifically, we utilized Hierarchical DBSCAN (HDBSCAN) [21,28], which is an extension of DBSCAN that finds clusters of varying densities by converting DBSCAN into a hierarchical clustering algorithm for clustering [29].

This choice was made to ensure that our clustering technique could handle data with diverse density distributions, resulting in more robust and accurate topic extraction from the input embeddings.

3.2.3. Topic Representation

The final step of the BERTopic process is the extraction of topics from the previously identified clusters. This step employs a novel variant of TF-IDF, known as concept-based TF-IDF (c-TF-IDF), to determine each cluster’s thematic essence. Unlike traditional TF-IDF, which evaluates word significance based on frequency across documents, c-TF-IDF extends this analysis to the corpus level, which enhances the contextual relevance of terms. Specifically, c-TF-IDF assesses the importance of terms not only based on their frequency in a specific topic, but also on their distribution across the entire corpus. This means that terms occurring frequently within a specific topic but infrequently in the rest of the corpus are assigned higher weights since they are considered to be more representative of the topic.

For a term x within class c ,

$$W_{x,c} = |tf_{x,c}| \times \log\left(1 + \frac{A}{f_x}\right) \quad (1)$$

where the terms have the following meanings:

$tf_{x,c}$: frequency of word x in class c .

f_x : frequency of word x across all classes.
 A : average number of words per class.

At this stage, we can gather information on the various topics, including the key terms associated with each topic. We can also visualize these topics and their key terms in various ways, including bar charts and heat maps, to better understand the distribution of topics and their relationships with each other.

Table 2 outlines the choices available at each stage of the BERTopic process, illustrating the flexibility and adaptability of this tool in addressing diverse analytical needs.

Table 2. The algorithms used for each stage of BERTopic.

Stage	Algorithms	Parameters
Embeddings	AraBert-v02	_____
Dimensionality reduction	UMAP	n_neighbors = 15 – n_components = 5 – metric = cosine
Clustering	HDBSCAN	default parameters
Topic representation	c-TF-IDF	bm25_weighting – reduce_frequent_words

4. Results and Discussion

A combination of quantitative and qualitative analyses was used to evaluate the results of the topic modeling. The quantitative assessment assumes independent functionality at each stage, allowing a thorough comparison of performance across a range of parameters. Simultaneously, a qualitative examination was carried out to evaluate the quality and relevance of each keyword in the generated topics. This evaluation focused on their alignment with topics classified by humans during a brainstorming session. The purpose of this qualitative analysis was to gain insights into the interpretability and effectiveness of the generated topics in capturing the essence of human-defined thematic categories.

4.1. Evaluation Metrics

The coherence scores serve as the most suitable metric when the topic model’s output is utilized by human users [30]. Various coherence scores are available, including *c_v* and *u_mass*. The *c_v* score is a prominent coherence metric. It constructs content vectors for words using co-occurrences and calculates the scores using cosine similarity and Normalized Pointwise Mutual Information (NPMI). The *u_mass* score evaluates the frequency of two words appearing together in the corpus. The overall coherence of the topic is established by averaging the pairwise coherence scores of the top N words describing the topic. Generally, the higher the coherence score is, the better the topics are [31]. In this work, we used both the *c_v* and *u_mass* to evaluate the quality of topics generated for each model for comparison validation and purposes. Their implementation was performed using the Gensim library (<https://radimrehurek.com/gensim/>, accessed on 15 April 2024).

4.2. Quantitative Evaluation

The LDA model was trained using various parameter configurations, including different numbers of topics (5, 10, 20, and 30) and maximum iteration counts of 100. The coherence ratings of all topics generated during the iterative process were recorded for each configuration. In contrast to LDA, BERTopic autonomously determines the optimal number of topics. Each resulting topic is defined by a unique set of top keywords, as demonstrated by a sample of five topics presented in Table 3. The annotations offer a concise understanding of each topic.

Table 3. Top 10 Arabic keywords and their corresponding English translations from five topics generated by BERTopic.

No.	Top-10 Keywords	Annotations
Topic 1	تأثير، سلبي، ضياع، للطلاب، تطوير، الزمن، النجاح، المهارات، الأكاديمي Negative, impact, loss, for students, development, time, success, skills, academic, waste	التحديات التعليمية Educational challenges
Topic 2	لتحفيزهم، يسبب، إخراج، تفادق، التلاميذ، لتحسين، الالتزام، للمواد To motivate, cause, direct, aggravate, students, to improve, commitment, to academic, subjects, lessons	تحفيز التلاميذ Motivating students
Topic 3	توفير، تعليمية، المنزلية، برامج، للطلاب، تنظيم، كيفية، انشاء، تشجيع، الطلاب Providing, educational, home, programs, for students, organizing, how to, create, encouraging, students	الدعم التعليمي المنزلي Home educational support
Topic 4	التعليمية، الفرص، زيادة، الفرصة، تجاهل، تفاقم، الأسرة، انقسام، فقدان، الانقسام Educational, opportunity, increase, opportunity, neglect, aggravation, family, division, loss, division	فرص التعليم Education opportunities
Topic 5	لتكوين، غياب، الأسرة، فاشل، جهل، المدرسي، مواطن، بالمساطر، ضعف، الأناية Formation, absence, family, failure, ignorance, school, citizen, rulers, weakness, selfishness	نظام التعليم Educational system

The results in Table 4 present a comparison between the performance of BERTopic, which is specifically configured for Arabic text, various configurations of Latent Dirichlet Allocation (LDA), and non-negative matrix factorization (NMF) in terms of topic modeling. Higher results are formatted in bold. This comparison highlights variations in coherence scores based on the number of topics and the presence of data preprocessing (DP). Performance was measured as specified earlier using two coherence scores: c_v and u_{mass} . Specifically, the BERTopic model configured for 29 topics achieved a c_v coherence score of 0.029 without DP, demonstrating its ability to generate coherent topics. However, when DP was applied, the c_v score decreased slightly to -0.002 . The c_v considers word similarity within the topic, which indicates the potential of BERTopic in capturing semantic relationships.

Similarly, the u_{mass} coherence score for BERTopic without DP was -6.630 , indicating a reasonable level of coherence, whereas with DP, the score decreased to -7.983 .

Comparatively, LDA models with DP exhibited similar trends, with higher numbers of topics generally resulting in lower coherence scores. For instance, LDA with 30 topics and DP achieved a c_v score of -0.180 and a u_{mass} score of -9.441 .

For NMF, the model with 30 topics achieved a c_v score of -0.018 without DP and 0.005 with DP. The u_{mass} coherence scores for NMF were -8.002 without DP and -8.689 with DP. Notably, the u_{mass} coherence for LDA configurations with a lower number of topics tended to be higher, indicating better performance on this metric. These findings suggest that both the number of topics and the application of DP significantly influence the coherence of topic modeling outcomes. The results indicate that while BERTopic shows promise in generating coherent topics, particularly without DP, NMF also demonstrates competitive performance, especially when DP is applied.

Table 4. Comparison between topic modeling models.

Topic Model	Number of Topics		Coherence (c_v)		Coherence (c_umass)	
	without DP	with DP	without DP	with DP	without DP	with DP
BERTopic	29	25	0.029	−0.002	−6.630	−7.983
LDA 5-100	5		−0.147	−0.151	−7.380	−7.240
LDA 10-100	10		−0.139	−0.171	−7.297	−8.292
LDA 20-100	20		−0.172	−0.169	−8.688	−8.623
LDA 30-100	30		−0.180	−0.195	−9.441	−9.454
NMF	30		−0.018	0.005	−8.002	−8.689

4.3. Qualitative Evaluation

Figure 4 presents a visual representation of the relationships between different topics, providing invaluable insight for brainstorming sessions. Participants can examine this graphical presentation to identify potential overlaps or thematic connections between topics, helping them to make decisions about merging topics to meet their specific brainstorming objectives.

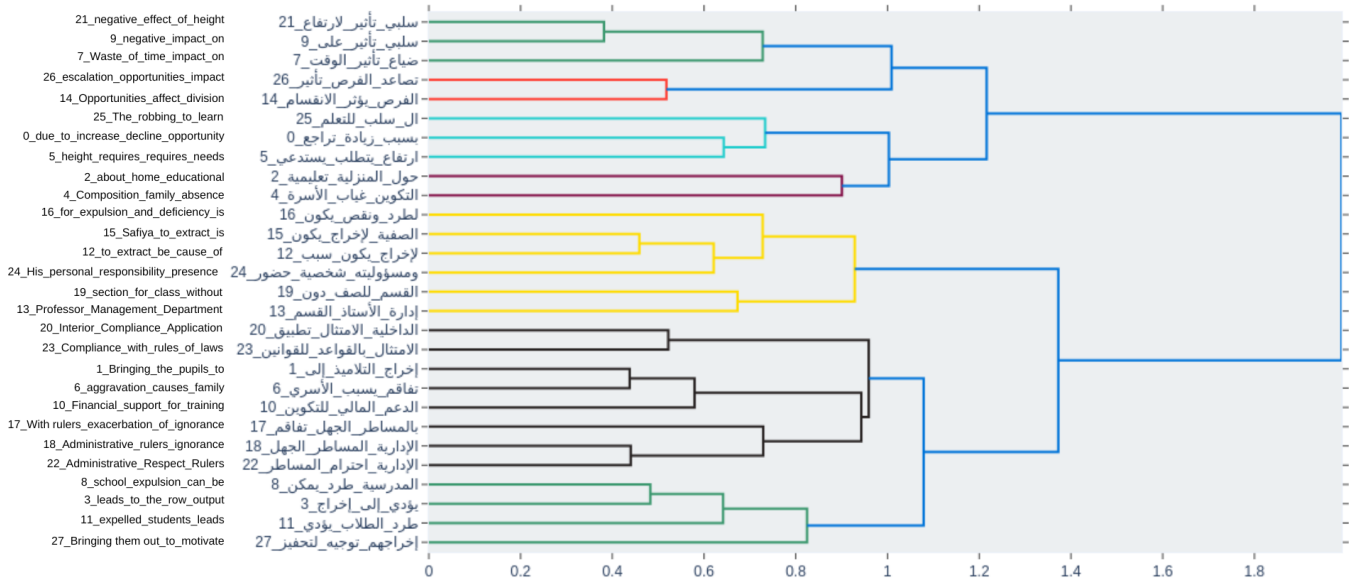


Figure 4. Hierarchical representation by BERTopic.

BERTopic has identified underlying topics within the analyzed text data, as demonstrated by the resulting topics. Figure 5 shows the c-TF-IDF scores of keywords within a sample of topics, highlighting the importance of each keyword in representing its respective topic. BERTopic’s approach to determining the number of generated topics is supported by its hierarchical reduction mechanism, which allows for the merging of topics based on their semantic similarities.



Figure 5. Topic words generated by BERTopic.

The interpretation of coherence scores, as demonstrated by the c_v and u_mass metrics, provides information about the quality of topic modeling results. While c_v coherence evaluates the semantic similarity of words within topics, u_mass coherence measures the degree of topic coherence using the relative distance between words in the document. Despite their importance, it is important to acknowledge their limits. For instance, c_v coherence may not completely reflect the semantic details of topics, while u_mass coherence may be impacted by document structure and length. As a result, it is critical to complement quantitative assessments with qualitative evaluations to obtain an extensive understanding of topic coherence and relevance.

On the other hand, data preprocessing impacts topic modeling outputs, influencing the quality and interpretability of generated topics. Tokenization and stopword removal are all preprocessing techniques that can have a considerable influence on coherence scores since they change the structure and content of the input text. While these techniques aim to improve the quality of textual data by standardizing and cleaning it, they may unexpectedly remove valuable information or introduce biases. Therefore, to obtain reliable and meaningful results, preprocessing techniques must be carefully considered, as well as their possible influence on topic modeling outcomes.

In comparison to traditional topic modeling approaches, BERTopic offers certain advantages in mitigating limitations. Its hierarchical reduction mechanism facilitates the merging of semantically similar topics, which may help to reduce the impact of over-splitting topics with only slightly different word choices. Furthermore, BERTopic’s interpretability through keywords with c_TF-IDF scores allows for human evaluation alongside quantitative metrics, providing a more comprehensive understanding of topic coherence.

This is a continuation of our previous work [4], where we explored the potential of optical recognition, spelling correction, and content-based information extraction. In this work, we used the BERTopic algorithm, which provides an efficient approach to automatically classifying ideas generated by as many participants as possible. The results of our evaluation, both quantitative and qualitative, demonstrate the superior performance of our system compared to traditional LDA models. The high coherence scores obtained with BERTopic confirm its ability to identify relevant thematic clusters in brainstorming data, providing a better understanding of the generated ideas. Additionally, the proposed system provides facilitators with an overview and visual tools to facilitate decision-making during brainstorming sessions by ensuring the in-depth analysis of topics and associated key terms. This automated approach improves the efficiency of brainstorming sessions by reducing the time needed to classify ideas while making it easier to identify key topics, helping facilitators in their decision-making process.

Figure 6 illustrates the brainstorming process within our model. Participants generated a multitude of ideas, which were recorded on Post-it notes. These notes were digitized using an optical character recognition (OCR) system to convert handwritten notes into machine-readable text. Our model helped to group related ideas under key themes, automatically categorizing them and providing facilitators with clear and actionable insights. This process significantly reduced the manual effort typically required to sort and analyze brainstorming outputs, proving particularly useful in managing the complexity of brainstorming sessions.

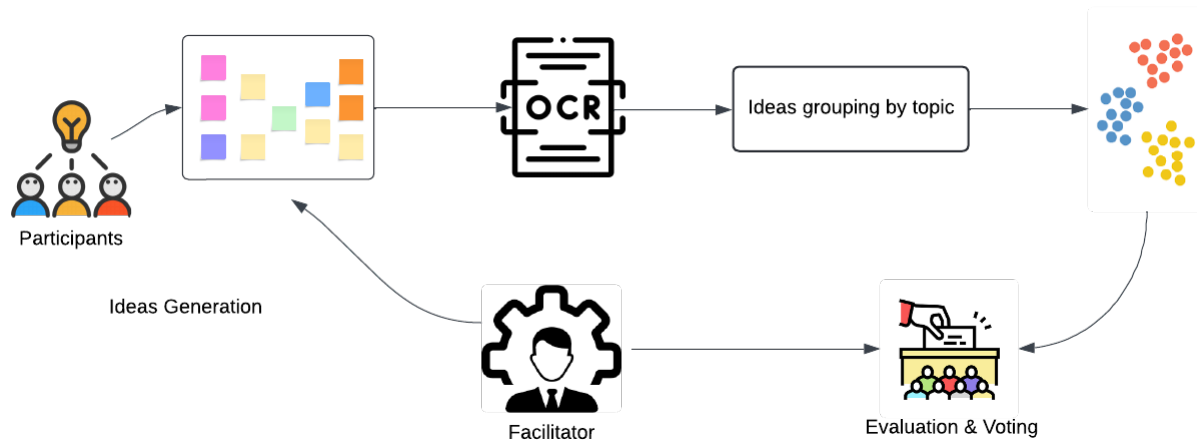


Figure 6. The brainstorming process within our model.

This work functions as a decision-support tool for the brainstorming session facilitator. Future research directions aim to enhance this system by incorporating continuous learning capabilities. By integrating the facilitator's adjustments and regulations made during the brainstorming session, the system can iteratively improve its topic modeling accuracy and effectiveness over time. Furthermore, future research will investigate the integration of qualitative evaluation methods, such as human expert reviews, alongside coherence metrics to provide a more nuanced understanding of topic relevance and coherence.

5. Conclusions

This study introduces an approach to automated brainstorming systems that utilizes the BERTopic algorithm to assist facilitators during brainstorming sessions. This system represents a significant step forward in the management of large volumes of ideas generated by diverse participants. The quantitative and qualitative evaluations demonstrate BERTopic's superior performance compared to traditional LDA models, resulting in high topic coherence and a clearer understanding of the generated ideas. Furthermore, the system provides facilitators with visual tools and an overview, enabling them to analyze topics and associated key terms in depth. This facilitates informed decision-making during brainstorming sessions. By automating idea classification and highlighting key themes, our approach demonstrably improves the efficiency of brainstorming sessions, reducing the analysis time and aiding facilitators in their decision-making process. As a significant contribution to the literature on automated brainstorming techniques, our system paves the way for future research and applications in the field of artificial intelligence to support creativity and innovation processes in organizations. Our study has several limitations. Firstly, it could benefit from incorporating user feedback to enhance the models and capture the nuances and intricacies of user preferences and perceptions. Secondly, the study does not account for recent advancements in language model understanding (LLM), as the research commenced prior to 2022. Hence, it is still likely that the models do not reflect the latest developments in LLM technology and methodology. Finally, it may be worth noting that the treatment of outliers in the data could present a challenge, as it has the potential to affect the accuracy and reliability of the results. These limitations highlight the importance of future research to address these aspects for more comprehensive and robust model development and evaluation.

Author Contributions: Conceptualization, M.E.H., Y.E.-S., A.C., and T.A.B.; methodology, M.E.H., Y.E.-S., and M.B. ; software, T.A.B. and A.C. ; validation, M.E.H., Y.E.-S., and M.B.; formal analysis, M.E.H., Y.E.-S., and M.B.; investigation, A.C.; data curation, A.C.; writing—original draft preparation, A.C. and T.A.B.; writing—review and editing, Y.E.-S. and M.E.H. and M.B.; supervision, Y.E.-S., M.E.H., and M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The code and data used in this work can be requested from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CBIE	Content-based information extraction
DP	Data preprocessing
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
HDBSCAN	Hierarchical DBSCAN
LDA	Latent Dirichlet Allocation
LSA	Latent Semantic Analysis
NLP	Natural language processing
NMF	Non-negative matrix factorization
OCR	Optical character recognition
pLSA	Probabilistic Latent Semantic Analysis
UMAP	Uniform Manifold Approximation and Projection

References

1. Memmert, L.; Tavanapour, N. Towards human-AI-collaboration in brainstorming: empirical insights into the perception of working with a generative AI. *ECIS 2023 Research Papers* https://aisel.aisnet.org/ecis2023_rp/219
2. Tang, L.; Peng, Y.; Wang, Y.; Ding, Y.; Durrett, G.; Rousseau, J.F. Less Likely Brainstorming: Using Language Models to Generate Alternative Hypotheses. *Proc. Conf. Assoc. Comput. Linguist. Meet.* **2023**, *2023*, 12532–12555. <https://doi.org/10.18653/v1/2023.findings-acl.794>.
3. Barki, H.; Pinsonneault, A. Small Group Brainstorming and Idea Quality: Is Electronic Brainstorming the Most Effective Approach? *Small Group Res.* **2001**, *32*, 158–205. <https://doi.org/10.1177/104649640103200203>.
4. Cheddak, A.; Ait Baha, T.; El Hajji, M.; Es-Saady, Y. Towards a Support System for Brainstorming Based Content-Based Information Extraction and Machine Learning. In *Proceedings of the International Conference on Business Intelligence, Beni-Mellal, Morocco, 27–29 May 2021*; Fakir, M., Baslam, M., El Ayachi, R., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 43–55. https://doi.org/10.1007/978-3-030-76508-8_4.
5. Paulus, P.B.; Baruah, J.; Kenworthy, J. Chapter 24—Brainstorming: How to get the best ideas out of the “group brain” for organizational creativity. In *Handbook of Organizational Creativity*, 2nd ed.; Reiter-Palmon, R., Hunter, S., Eds.; Academic Press: Cambridge, MA, USA, 2023; pp. 373–389. <https://doi.org/10.1016/B978-0-323-91840-4.00019-0>.
6. Russell, T.M. Interactive Ideation: Online Team-Based Idea Generation Versus Traditional Brainstorming. Ph.D. Thesis, University of Minnesota, Minneapolis, MN, USA, 2019.
7. Paulus, P.B.; Kenworthy, J.B. Effective brainstorming. In *The Oxford Handbook of Group Creativity and Innovation*; Oxford University Press: Oxford, UK, 2019; pp. 287–305.
8. Paulus, P.B.; Yang, H.C. Idea generation in groups: A basis for creativity in organizations. *Organ. Behav. Hum. Decis. Process.* **2000**, *82*, 76–87.
9. Deckert, C.; Mohya, A.; Suntharalingam, S. Virtual whiteboards & digital post-its—incorporating internet-based tools for ideation into engineering courses. In *Proceedings of the SEFI 2021: 49th Annual Conference Blended, Virtual, 13–16 September 2021*; pp. 1370–1375.
10. Dhaundiyal, D.; Pant, R. *Tools for Virtual Brainstorming & Co-Creation: A Comparative Study of Collaborative Online Learning*; Indiana University Southeast: New Albany, IN, USA, 2022.
11. Wieland, B.; de Wit, J.; de Rooij, A. Electronic Brainstorming With a Chatbot Partner: A Good Idea Due to Increased Productivity and Idea Diversity. *Front. Artif. Intell.* **2022**, *5*, 880673. <https://doi.org/10.3389/frai.2022.880673>.

12. Ekramipooya, A.; Boroushaki, M.; Rashtchian, D. Application of natural language processing and machine learning in prediction of deviations in the HAZOP study worksheet: A comparison of classifiers. *Process Saf. Environ. Prot.* **2023**, *176*, 65–73. <https://doi.org/10.1016/j.psep.2023.06.004>.
13. Evangelopoulos, N.E. Latent semantic analysis. *Wiley Interdiscip. Rev. Cogn. Sci.* **2013**, *4*, 683–692. <https://doi.org/10.1002/wcs.1254>.
14. Hofmann, T. Unsupervised Learning by Probabilistic Latent Semantic Analysis. *Mach. Learn.* **2001**, *42*, 177–196.
15. Zhang, C.; Yang, Q.; Zhang, J.; Gou, L.; Fan, H. Topic Mining and Future Trend Exploration in Digital Economy Research. *Information* **2023**, *14*, 432.
16. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
17. Hwang, S.J.; Lee, Y.K.; Kim, J.D.; Park, C.Y.; Kim, Y.S. Topic Modeling for Analyzing Topic Manipulation Skills. *Information* **2021**, *12*, 359.
18. Egger, R.; Yu, J. A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts. *Front. Sociol.* **2022**, *7*, 886498. <https://doi.org/10.3389/fsoc.2022.886498>.
19. Mendonça, M.; Figueira, Á. Topic Extraction: BERTopic’s Insight into the 117th Congress’s Twittersverse. *Informatics* **2024**, *11*, 8. <https://doi.org/10.3390/informatics11010008>.
20. Crețulescu, R.G.; Morariu, D.I.; Breazu, M.; Volovici, D. DBSCAN algorithm for document clustering. *Int. J. Adv. Stat. It C Econ. Life Sci.* **2019**, *9*, 58–66.
21. Ros, F.; Guillaume, S.; Riad, R.; El Hajji, M. Detection of natural clusters via S-DBSCAN a Self-tuning version of DBSCAN. *Knowl.-Based Syst.* **2022**, *241*, 108288. <https://doi.org/10.1016/j.knosys.2022.108288>.
22. Barman, P.C.; Iqbal, N.; Lee, S.Y. Non-negative Matrix Factorization Based Text Mining: Feature Extraction and Classification. In *International Conference on Neural Information Processing*; King, I., Wang, J., Chan, L.W., Wang, D., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; pp. 703–712.
23. de Groot, M.; Aliannejadi, M.; Haas, M.R. Experiments on Generalizability of BERTopic on Multi-Domain Short Text. *arXiv* **2022**, arXiv:2212.08459. <https://doi.org/10.48550/arXiv.2212.08459>.
24. Shorten, C.; Khoshgoftaar, T.M.; Furht, B. Text Data Augmentation for Deep Learning. *J. Big Data* **2021**, *8*, 101. <https://doi.org/10.1186/s40537-021-00492-0>.
25. Reimers, N.; Gurevych, I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *arXiv* **2019**, arXiv:1908.10084. <https://doi.org/10.48550/arXiv.1908.10084>.
26. McInnes, L.; Healy, J.; Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv* **2018**, arXiv:1802.03426.
27. Grootendorst, Maarten BERTopic: Neural topic modeling with a class-based TF-IDF procedure *arXiv* **2022**, arXiv:2203.05794.
28. Bhattacharjee, P.; Mitra, P. A survey of density based clustering algorithms. *Front. Comput. Sci.* **2021**, *15*, 1–27.
29. Malzer, C.; Baum, M. A Hybrid Approach To Hierarchical Density-based Cluster Selection. In *Proceedings of the 2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, Karlsruhe, Germany, 14–16 September 2020; pp. 223–228. <https://doi.org/10.1109/MFI49285.2020.9235263>.
30. Abdelrazek, A.; Eid, Y.; Gawish, E.; Medhat, W.; Hassan, A. Topic modeling algorithms and applications: A survey. *Inf. Syst.* **2023**, *112*, 102131. <https://doi.org/10.1016/j.is.2022.102131>.
31. Röder, M.; Both, A.; Hinneburg, A. Exploring the Space of Topic Coherence Measures. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining. Association for Computing Machinery, WSDM ’15*, Shanghai, China, 2–6 February 2015; pp. 399–408. <https://doi.org/10.1145/2684822.2685324>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.