IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 13, May 2026
Impact Factor: 8.2

Extractive Dialogue Summarization with Keyword Extraction and Sentence Ranking

Muthulakshmi.A[1], Kanimozhi.V[2], Varsha.S[3]
Department of AI&DS[1-3]
MEPCO Schlenk Engineering College, Sivakasi, India

Abstract: *Automatic text summarization is a foundational task in natural language processing, aimed at condensing vital information while maintaining the fundamental essence of a provided text. Traditional text summarization methods may fall short when applied to dialogue summarization, which specifically aims to capture the essence of conversations. In this paper, we introduce a specialized method designed for dialogue summarization, employing techniques such as keyword extraction, BERT embeddings and Page Rank. The scoring system takes into account cosine similarity and adds extra points for sentences containing keywords. Unlike conventional summarization approaches tailored for continuous texts like articles and stories, our method addresses the unique challenges posed by conversational content. This strategy allows us to create straightforward and brief summaries that capture the fundamental elements of spoken dialogues between two individuals. Significantly, our approach is flexible and can be applied to any conversation involving two people. The effectiveness of our approach is evaluated using the Rouge score on any conversational dialogues, showcasing its capability to generate compelling and informative summaries. We carry out experiments to assess the efficacy of our proposed approach in dialogue summarization tasks. Results demonstrate a notable improvement over strong baseline methods, with our approach outperforming by 8 percentage points on ROUGE-1 and 5 percentage points on ROUGE-2. These findings underscore the effectiveness of our model in dialogue summarization.*

Keywords: Dialogue Summarization, KeyBERT, Page Rank, Topic Modeling

I. INTRODUCTION

In recent years, the proliferation of multi-sourced information on the Internet has created a vast reservoir of textual content accessible to the public. Effectively summarizing this abundance of text has become increasingly essential, as manual summarization by humans becomes impractical due to the sheer volume of information. Automatic text summarization has arisen as a valuable tool in the information age, capable of generating concise, fluent, and coherent summaries while retaining the original content and meaning. This capability extends across diverse fields such as news reports, electronic correspondences, telephone logs, medical files, and text communications on mobile devices. Several online summarization tools, including Microsoft News2[40], Google1, MEAD[37], SWESUM[38], and WikiSummarizer[39], cater to specific domains. Automatic text summarization methods broadly fall into two categories: extractive, which selects and reproduces key sentences verbatim, and abstractive, which generates new sentences.

Our paper focuses on summarizing conversations, particularly spoken dialogues. However, the automatic summarization of spoken dialogue conversations poses unique challenges:

- They are not uninterrupted passages but rather encompass exchanges between two individuals.
- They are frequently lengthy and interspersed with casual conversation, comprising numerous sentences that may lack relevance or coherence.
- They contain numerous poorly constructed, grammatically incorrect sentences.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 13, May 2026 Impact Factor: 8.2

- They may lack proper punctuation or have incorrect punctuation based on the conversational pauses perceived by annotators, rendering them difficult to read.
- Current open-source summarization tools do not excel in summarizing spoken dialogues.

Our focus lies in summarizing conversations, specifically spoken dialogues. Existing tools may not fully meet the requirements for summarizing these dynamic and unstructured discussions. To handle this, we propose an innovative method that incorporates advanced techniques. This involves employing topic modeling and KeyBERT for keyword extraction, along with sentence selection based on scores calculated using BERT embeddings and PageRank. The scoring system takes into account cosine similarity and adds extra points for sentences containing keywords. This approach enables us to craft clear and concise summaries, capturing the essence of conversations between two individuals in spoken dialogues. Importantly, this method is versatile and applicable to any dialogue conversation involving two persons.

To overcome the challenges posed by spoken dialogue conversations, we propose a multifaceted approach employing advanced techniques:

- Implement algorithms for topic modeling to identify key themes or subjects within the spoken dialogue, aiding in comprehending the main ideas discussed during the conversation.
- Utilize KeyBERT for extracting crucial keywords from the dialogue, identifying significant terms that encapsulate core topics and concerns.
- Apply BERT Embedding to represent the semantics of sentences in the dialogue. Utilize cosine similarity to measure the similarity between sentences, considering their contextual meanings.
- Implement PageRank, a link analysis algorithm, to assign importance scores to sentences based on their relationships and significance within the dialogue. This aids in selecting pivotal sentences for inclusion in the summary.

II. RELATED WORKS

P.k. Biswas et al. [1] introduces a pioneering method that combines topic modeling and sentence selection with punctuation restoration to address unique challenges, such as conversational nature and multiple topics within calls. By leveraging these techniques, the proposed approach generates concise and readable call summaries, marking a significant advancement in the field. B. Ma et al. [2] introduces leveraging Distantly Supervised Machine Reading Comprehension in customer service. This approach eliminates the dependency on annotated data, utilizing distant supervision and speaker role information to enhance efficiency and accuracy in dialogue summarization for improved customer service. H. Lin et al. [3] introduces the novel task of Topic-Oriented Dialogue Summarization (TODS), addressing challenges in learning semantic information, locating topic related content, and distinguishing summaries for different topics. Through three proposed auxiliary tasks, the paper establishes their effectiveness in generating high-quality summaries and emphasizes their significance in advancing research in the field of TODS. Y. Liu et al. [4] presents a comprehensive framework employing pretrained BERT models for extractive and abstractive summarization, achieving state-of-the-art results. The authors emphasize the significance of document-level encoding, laying the groundwork for advancements in text summarization techniques across diverse datasets and evaluation protocols. J. Shin et al. [5] introduces a novel approach to enhance factual consistency in abstractive summaries through token level correction. By employing a sequential model with a summary fact checker, token fact checker, and fact, the authors address factual inconsistencies more comprehensively, emphasizing the importance of accurate fact correction in real-world applications. H. Zhu et al. [6] introduces WIKI REF, a large query-focused summarization dataset constructed from Wikipedia, and proposes the Q-BERT model. Leveraging data augmentation and fine-tuning, the model demonstrates enhanced performance on query-focused summarization, emphasizing the crucial role of augmented data and quality in improving data-driven summarization models. Alexandra N. Uma et al. [7] presents an experimental com parison of extractive summarization models for callcenter dialogue without relying on gold summaries for training. It highlights the effectiveness of models considering word importance, recognizes the efficacy of simple baselines like

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 13, May 2026 Impact Factor: 8.2

Lead-7, and advocates for objective and subjective evaluations in assessing summarization model suitability, suggesting future research directions in larger annotated call datasets and supervised extractive summarization techniques trained on dialogue datasets.

Rubayyi Alghamdi et al. [8] surveys the significance of topic modeling in text mining, emphasizing its role in analyzing unstructured text data for pattern discovery, document connectivity, and representation as topic mixtures. It underscores the relevance of incorporating time in topic evolution models and offers insights into diverse topic modeling methods and their applications in text analysis. Z. Wu et al. [9] focuses on a novel approach to summarizing lengthy documents through the extraction of knowledge graphs. It likely delves into methodologies for identifying and representing key information in the form of a knowledge graph, summarizing complex content more effectively. The authors may explore techniques for handling lengthy texts, highlighting challenges and proposing solutions M. Zhong et al. [10] likely introduces an approach that frames.

extractive summarization as a text matching problem. The authors likely propose a methodology where relevant sentences are identified based on their similarity to a predefined summary or key information. This approach may leverage text matching algorithms and techniques to determine the importance of sentences within the document. V. Gupta et al. [11] conducts a thorough review of various extractive summarization techniques, encompassing statistical, graph-based, and machine learning algorithms, while emphasizing their strengths, limitations, and real-world applications. It addresses challenges, such as redundancy and coherence preservation, and provides valuable insights into recent advancements and future directions, making it a comprehensive resource for researchers and practitioners in the field. S. Verma et al. [12] proposes a text summarization approach using a deep learning model that consists of three phases: feature extraction, feature enhancement, and summary generation. The approach aims to improve the selection of sentences for the summary and uses a Restricted Boltzmann Machine to enhance and abstract features. The effectiveness of the approach is demonstrated through experimentation on several articles. D. Miller et al. [13] introduces a novel approach using BERT for lecture content summarization. The model captures linguistic nuances, improves quality, and can handle lecture-based content, including technical terminology and diverse speaking styles. J. Devlin et al. [14] introduces a ground-breaking language model, BERT, designed for pretraining deep bidirectional transformers. It enables learning complex language patterns and is transferable to various tasks, including question answering and text classification. Y. Liu et al. [15] introduces a fine-tuning approach to BERT, enhancing information retrieval, generating coherent summaries, and improving content condensation. The paper aims to showcase the model's adaptability and utility in generating high-quality summaries from diverse textual inputs. S. Narayan et al. [16] explores extractive summarization, focusing on enhancing sentence selection through reinforcement learning. Traditional methods like graph algorithms and deep learning models such as RNNs and BERT embeddings are examined. Contextual understanding via neural networks is emphasized for capturing sentence relationships. Reinforcement learning iteratively refines rankings based on rewards, optimizing summarization. Comparative analyses ad dress challenges like data sparsity and diverse evaluation metrics. Overall, the paper advances sentence ranking in summarization, highlighting reinforcement learning's role in producing coherent summaries. Z. Wu et al. [17] introduces a novel method for summarizing lengthy documents using knowledge graphs. It likely discusses techniques for identifying and representing key information effectively. Natural language processing and machine learning are probably ap plied for information extraction and structuring. Emphasis is placed on improving accessibility and comprehension through summarization. Graph-based representation is integrated for a condensed yet comprehensive overview. Potential applications include information retrieval and question-answering systems. N. Franciscus et al. [18] likely explores dependency graphs for summarizing short texts effectively. It discusses leveraging syntactic relationships to identify key content. Natural language processing aids in constructing and analyzing these graphs. Emphasis is on capturing essential meaning and preserving context. Practical applications include social media analytics and sentiment analysis. Overall, it aims to understand the potential and challenges of dependency graphs in short text summarization. Q. Zhou et al. [19] proposes a neural network-based approach for document summarization. It employs a joint learning framework to simultaneously score and select sentences. This approach captures semantic

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 13, May 2026 Impact Factor: 8.2

relationships and contextual information effectively. Dual-objective learning enables informed sentence selection using local and global context. Reinforcement learning further fine-tunes the model, improving summarization quality. Experimental results demonstrate significant improvements over traditional methods, advancing neural-based summarization techniques. J. Cheng et al. [20] introduces a novel neural summarization approach, extracting both sentences and words for concise summaries. Dual extraction dynamically selects key sentences and essential words for flexibility. An attention-based mechanism assigns importance scores to sentences and identifies crucial words within them. Reinforcement learning tackles abstractive summarization, optimizing summary quality. Experimental results demonstrate improved content retention and informativeness.

Guo's work advances neural summarization with a balanced, coherent approach, promising high-quality, content-preserving summaries through dual extraction. K.Jezek et al. [21] explores automatic text summarization, covering both extractive and abstractive techniques. It dis cusses advancements in NLP and machine learning, addressing challenges of coherence and relevance. Evaluation metrics for summarization quality are likely discussed, along with the impact of deep learning models. Significance in domains like information retrieval and document clustering is explored. Overall, it aims to offer a comprehensive overview for re searchers and practitioners in automatic text summarization. J. Neto et al. [22] likely introduces a machine learning-based method for automatic text summarization. It explores various algorithms and models to extract key information. Advantages include adaptability to diverse text types and improved accuracy. The training process involves identifying and prioritizing essential content, addressing challenges like diverse language structures. Features like sentiment analysis or topic modeling may enhance summarization quality. Practical applications in information retrieval and document management are likely explored. Overall, the paper aims to demonstrate the effective ness of machine learning in producing high-quality summaries from diverse textual inputs. K. Kaikhah et al. [23] likely focuses on leveraging neural network architectures to enhance automatic text summarization. It explores models like RNNs or transformer-based architectures such as BERT. Advantages include capturing complex linguistic patterns and context. Training involves optimizing networks to learn textual relationships for effective summarization. Challenges like coherence and relevance are likely addressed. Pre-trained language models contribute to efficiency. Practical applications include improving search engine results and information retrieval. Overall, the goal is to demonstrate neu ral networks' effectiveness in automating summarization and producing coherent summaries from diverse textual sources. L. Suanmali et al. [24] likely proposes integrating fuzzy logic into text summarization to enhance accuracy and effectiveness, leveraging its ability to handle uncertainty and imprecision. A fuzzy inference system may be introduced to capture nuanced linguistic relationships. Methodology likely involves considering degrees of importance or relevance in sentence extraction. Fuzzy logic offers solutions to challenges like subjective language interpretation and context-dependency. Training and optimization of the system for various texts are discussed. Practical implications include improved summary coherence and adaptability to diverse linguistic styles. Overall, the goal is to demonstrate how fuzzy logic refines summarization, providing nuanced and context-aware summaries.

R. Nallapati et al. [25] introduces SummaRuNNer, leveraging recurrent neural networks for text summarization. Sum maRuNNer focuses on extracting salient information from documents. Its architecture likely involves processing sentence sequences to rank key information. Training on large datasets enables pattern learning for coherent summaries. Practical applications across domains, including document summarization, are discussed. Overall, SummaRuNNer aims to advance extractive summarization by enhancing crucial information extraction. J. Xu et al. [26] introduces a novel extractive summarization approach incorporating neural networks and syntactic compression. The model likely identifies key information and applies compression to enhance summary quality and conciseness. Neural architecture details may explain how syntactic structures are processed and compressed. Training likely involves learning optimal compression strategies for effective distillation. Practical applications include improved document understanding and enhanced information retrieval. The paper aims to showcase the efficacy of combining neural networks with syntactic compression, presenting a sophisticated method for producing concise and contextually relevant summaries. Q. Zhou et al. [27] introduces a unified framework for neural extractive summarization, addressing coherence and informativeness challenges. It integrates a neural network based sentence scoring mechanism to capture

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 13, May 2026
Impact Factor: 8.2

contextual relationships[36]. Simultaneously, a selection mechanism optimizes sentence extraction for summaries. The joint consideration enhances decision-making iteratively. The framework outperforms baseline methods, generating coherent and content rich summaries. It contributes a holistic solution to scoring and selection intricacies, advancing automatic summarization systems. Overall, it provides valuable insights for improving efficiency and quality in summarization. X. Zhou et al. [28] introduces a unified framework for neural extractive summarization, addressing coherence and informativeness challenges. It integrates a neural network-based sentence scoring mechanism to capture contextual relationships. Simultaneously, a selection mechanism optimizes sentence extraction for summaries. The joint consideration enhances decision-making iteratively. The framework outperforms baseline methods, generating coherent and content-rich summaries. It contributes a holistic solution to scoring and selection intricacies, advancing automatic summarization systems. Overall, it provides valuable insights for improving efficiency and quality in summarization. D. Wang et al. [29] introduces a novel approach to document summarization using heterogeneous graph neural networks (HGNN). It models complex relationships among document entities through a heterogeneous graph structure. HGNN captures semantic interactions between diverse nodes, aiding comprehensive content understanding. The approach exploits intra sentence and inter-sentence relationships for informed sentence ranking. Advantages include encoding diverse contextual information. Experimental results demonstrate HGNN's superior performance in generating high-quality summaries. Its application marks a significant advancement in extractive summarization, merging graph-based methods with natural language processing for future research exploration. J. Xu et al. [30] introduces a novel approach to extractive text summarization by incorporating discourse information. It enhances coherence and informativeness by considering discourse structure. A neural network-based model integrates discourse-aware features, capturing relationships between sentences. Discourse markers and relations are incorporated to understand contextual flow. Comprehensive experiments demonstrate superior performance in coherence and informativeness. Leveraging discourse information is crucial for generating contextually relevant summaries, marking a significant advancement in extractive summarization techniques.

III. PROPOSED WORK

A. Model Overview

In our methodology, illustrated in Fig. 3, we commence with a robust pre-processing phase aimed at enhancing the quality of the input dialogue text. This initial step involves cleaning and refining the textual dialogue data. Subsequently, keywords are dynamically extracted using various methods, such as traditional topic modeling techniques like LDA[33], LSI[34], HDP[35], or KeyBERT model. BERT, recognized for its contextual understanding, is then utilized to embed the text. Employing cosine similarity, our model assesses semantic relationships across different sections of the embedding, thereby enriching the information retrieval process. To prioritize significant con tent, we make use of a PageRank algorithm that constructs a graph based representation of the text, assigning importance scores to sentences based on their relationships within the document. During the process, the presence of a keyword in a sentence contributes to an increased score for that sentence. This strategic addition ensures that sentences containing relevant keywords receive heightened importance within the graph based representation. Each sentence is meticulously assigned a score based on a combination of its PageRank score and the presence of keywords. Ultimately, our approach culminates in the generation of a coherent summary that effectively captures the essence of the input dialogue. This showcases a harmonious integration of pre-processing, keyword extraction, embedding, similarity analysis, graph-based ranking, and summary synthesis, underscoring the comprehensive nature of our methodology.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 13, May 2026
Impact Factor: 8.2

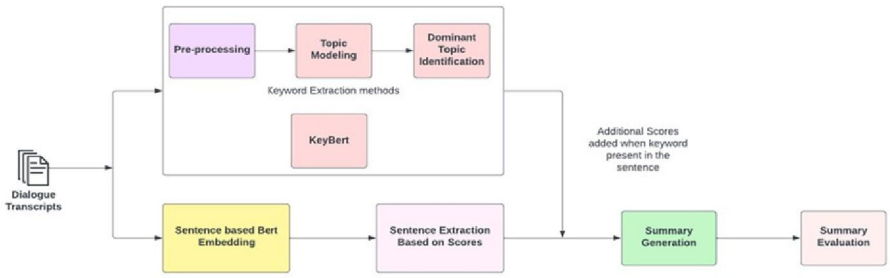

Figure 1 - System Design

B. Pre-processing

Our pre-processing pipeline employs various techniques:

1. *Tokenization:* Tokenization segments the text into individual words, enabling subsequent analysis to understand the linguistic structure of the dialogue.
2. *Contraction Mapping:* Contraction mapping resolves con tractions, such as converting ”don’t” to ”do not,” ensuring consistent and expanded representations for improved linguistic processing.
3. *Punctuation Removal:* Punctuation removal eliminates extraneous punctuation marks, streamlining the text for fo cused analysis while preserving the core information in the conversation.
4. *Stopword Removal and Custom Stopword Removal:* Stopword removal filters out common words, enhancing the significance of the remaining words, with an additional layer of custom stopwords specific to the dialogue context.
5. *Length Filtering:* Length filtering removes short words (length ≤ 4), streamlining the text by excluding less informative elements and focusing on meaningful content.
6. *Lowercasing:* Lowercasing standardizes the text to lowercase, ensuring consistency and simplifying subsequent pro cessing and analysis steps.

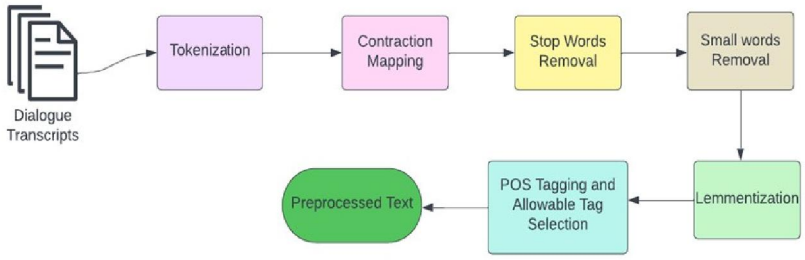

Figure 2 - Pre-processing Phases

C. Keyword Extraction

Extracting keywords is crucial for distilling important information from text. In our paper, we use advanced techniques to make this process more efficient for finding and summarizing information. We focus on two main ways of extracting keywords: 1) Topic Modeling 2) KeyBERT 1) Topic Modeling: In our methodology, [1] when specifying the topic model type (MT) during the procedure, we generate multiple topic models (TM) of the specified type using the entire conversation of the dialogue as a unified document. Utilizing documents (D), corpus (C), and vocabulary (V) from the corresponding transcripts, we systematically vary hyper-parameter values, such as the number of topics (K), within predefined ranges (e.g., 5-50) at predefined steps. The coherence scores (CS) for each model are computed, and the

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 13, May 2026 Impact Factor: 8.2

optimal topic model (TMmax) along with its associated hyper parameter values, resulting in the highest coherence scores, is identified using Equation (1)

$$\text{TMmax} = argmax_{TM}(\text{CS}(\text{TM}(\text{V}, \text{C}, \text{D}, \text{MT}, \text{K})) : \text{K} = \text{NumberOfTopics}, . . ,50) \quad (1)$$

If the topic model type is not specified, the process is conducted concurrently for three different topic model types (Latent Dirichlet Allocation (LDA), Latent Semantic Indexing (LSI), and Hierarchical Dirichlet Process (HDP)) using Equation (2). The algorithm identifies the topic models and associated hyper-parameter values that yield the best scores among the models of all three types for both participants.

$$TM_{max} = argmax_{\text{MT}\in\text{LDA,LSI,HDP}}(((\text{CS}(\text{TM}(\text{V,C,D,MT,K})):\text{K=NumberOfTopics,..,50)) \quad (2)$$

The algorithm relies extensively on the Python-based genism package for topic modeling. Dominant topic Identification: To identify the most prominent theme in each dialogue, we determine the dominant topic(DT) using the coherence scores derived from the previously generated topic models. Iterating through all transcripts, we locate the tuple with the highest coherence score within the list of topics. The associated topic ID from this tuple is then assigned to a variable, representing the dominant topic for that specific dialogue. The formula for this process is expressed as follows:

$$DT = argax_{topic}(\text{coherence\_score(topic)}: \text{topic} \in \text{list\_of\_T}M_{max}topics \quad (3)$$

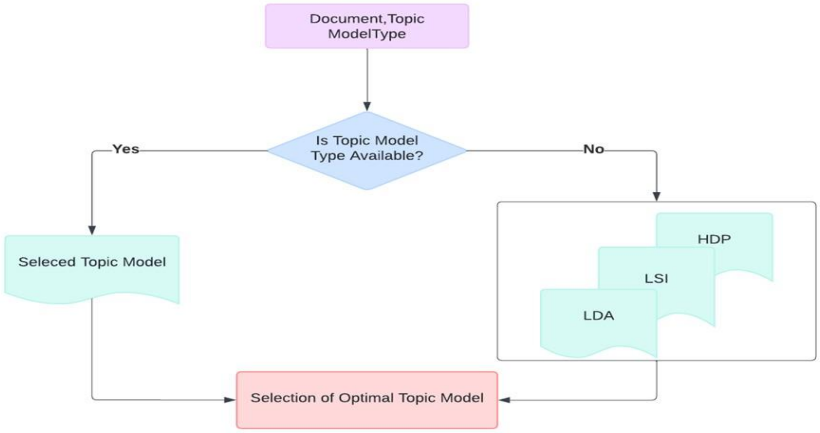

Figure 3 - Optimal Topic modelling

2) KeyBERT: KeyBERT is a transformer-based keyword extraction library designed to automatically identify and extract the most relevant keywords from a given text or document. It leverages transformer-based models, particularly BERT (Bidirectional Encoder Representations from Transformers)[13][14], to generate contextual embeddings for words in the text. The KeyBERT[32] model utilizes the 'all-mpnet-basev2' pre-trained transformer model for generating contextual embeddings. The extraction process involves excluding common English stop words, and the model identifies the top 10 keywords based on their relevance scores, providing a curated list of the most crucial terms within the text. This streamlined approach improves the efficiency of information retrieval and summarization.

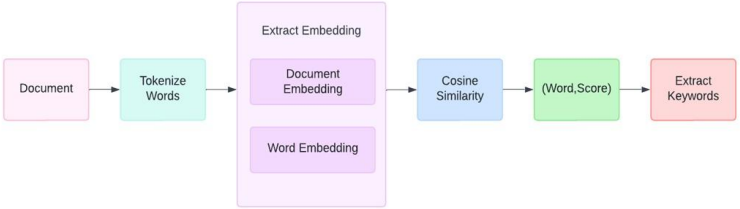

Figure 4 - KeyBERT design

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 13, May 2026
Impact Factor: 8.2

Bert Embedding

BERT embeddings, derived from the BERT model[13][14], capture contextualized word representations by conveying nuanced meanings within entire sentence contexts. Unlike traditional word embeddings, BERT embeddings are dynamic, context-dependent, and consider bidirectional relationships during training. The integral tokenization process in BERT involves breaking input text into subword tokens for improved vocabulary handling and effective capture of contextual relationships.

In the subsequent model inference stage, the pre-trained BERT model processes tokenized input, leveraging its transformer architecture and considering bidirectional dependencies through the attention mechanism. The resulting PyTorch tensor output is both flexible and efficient, encapsulating contextualized embeddings and contributing to computational efficiency in downstream tasks. Strategic use of 'torch.no grad()' during inference enhances efficiency, particularly in feature extraction. The culmination of BERT feature extraction involves obtaining the last hidden state, producing contextualized embeddings for each token that reflect semantic relationships and contextual nuances. The emphasized significance of the [CLS] token, strategically placed at the sequence's outset, enhances efficiency in downstream natural language processing tasks.

Score Based Sentence Ranking and Summary Generation

a) Cosine Similarity: Cosine similarity serves as a metric utilized to evaluate the likeness between two vectors within a multi-dimensional space. In the context of natural language processing and text analysis, cosine similarity is frequently utilized to assess the resemblance between the vector representations of two pieces of text. When utilizing BERT embeddings, each sentence or document is transformed into a high-dimensional vector representation. The cosine similarity between these vectors is then calculated to quantify the degree of similarity between the corresponding pieces of text. This similarity metric spans from -1 (entirely dissimilar) to 1 (completely similar), where 0 denotes no similarity. The cosine similarity (Cosine Similarity) between two vectors Emb1 and Emb2 can be expressed using the dot product (Emb1.Emb2) and the magnitudes (‖Emb1‖ and ‖Emb2‖) of the vectors as follows:

$$Cosind\ Similarity\ (Emb1, Emb2) = \frac{Emb1\ .\ Emb2}{||Emb1||\ |Emb2||} \quad (4)$$

Here,

• Emb1.Emb2 represents the dot product of vectors Emb1 and Emb2.

• ‖Emb1‖ and ‖Emb2‖ represent the magnitudes or Euclidean norms of vectors Emb1 and Emb2 respectively.

In the context of dialogue summarization, employing cosine similarity with BERT embeddings enables the algorithm to assess the semantic similarity between sentences or portions of dialogue. This measure is valuable for tasks such as sentence selection, where the goal is to identify sentences that capture similar or related information for inclusion in a concise summary.

b) Page Rank: PageRank, developed by Google, calculates a numerical significance for each element in a linked set of documents, such as web pages, to determine their importance. In the summarization context, PageRank is modified to gauge the significance of sentences within a document or dialogue. The fundamental concept is to regard sentences as vertices in a graph, and the connections between them as edges.

The process involves the following steps:

1. Graph Representation: Build a graph where sentences serve as vertices, and the connections between them represent edges. The intensity of the linkage between two sentences is determined by their semantic similarity, often measured using cosine similarity with BERT embeddings.

2. PageRank Computation: Apply the PageRank algorithm to calculate scores for each sentence based on the graph structure. This step identifies the sentences that are central or influential in the dialogue.

3. Keyword Integration: Integrate keyword information into the PageRank scores. Sentences containing relevant keywords are given an additional weight or bonus score, emphasizing their importance in conveying key information.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 13, May 2026
Impact Factor: 8.2

$$PageR(Emb) = \frac{1-df}{N} + df \sum_{e \in edges_{Emb}} (\frac{}{t}$$

Where:
• PageR(Emb) is the PageRank score for node Emb.
• df is the damping factor.
• N is the total number of vertices in the graph.
• edges Emb are the edges pointing to node Emb.
• total Edges is the total number of edges originating from the source node of edge.
• B is the additional bonus score if a keyword is present in the sentences.

c) Summary Generation: In summary generation, PageRank can be enhanced by incorporating keyword information to assign additional weight to sentences that contain relevant keywords. This involves computing the PageRank scores for sentences based on their relationships (connections) with other sentences, and then adding a bonus score or weight to sentences that include keywords obtained from the dialogue using techniques like topic modeling or keyword extraction algorithms. Select sentences with high combined scores (PageRank + Bonus for Keywords) for inclusion in the summary. This ensures that the summary not only captures the overall structure of the dialogue but also emphasizes sentences containing critical keywords, enhancing the informativeness of the generated summary.

By leveraging PageRank along with keyword-based bonus scores, the summarization process becomes more nuanced and contextually relevant. This approach helps prioritize sentences that are both structurally important and rich in relevant content, contributing to the generation of comprehensive and informative dialogue summaries.

Summary Evaluation

The effectiveness of our proposed summarization model is assessed using established evaluation metrics, specifically ROUGE and Bleu scores. These metrics serve as robust measures for the quality of generated summaries.

1) ROUGE Scores: ROUGE-1, ROUGE-2, ROUGE-L, ROUGE-F, ROUGE Precision, and ROUGE Recall are employed to evaluate the intersection between the produced summaries and original summaries.

a) ROUGE-1: Evaluates the intersection of unigrams (single words) between the produced and original summaries.

b) ROUGE-2: Evaluates the intersection of bigrams (two consecutive words) between the produced and original summaries.

c) ROUGE-L: Focuses on the Longest Common Subsequence (LCS) of words, providing a recall-oriented measure.

d) ROUGE-F (F1 Score): ROUGE-F, the harmonic mean of Precision and Recall, offers a balanced assessment, considering both False Positives (precision) and False Negatives (recall). It furnishes a comprehensive gauge of the model's capacity to strike a harmonious equilibrium between precision and recall in summary generation.

e) ROUGE Precision: ROUGE Precision specifically evaluates the precision of n-grams in the produced summary compared to the reference summary, evaluating the proportion of correctly generated n-grams.

f) ROUGE Recall: ROUGE Recall, which complements precision, assesses the model's ability to capture relevant information in the reference summary. It measures the proportion of relevant n-grams in the reference summary that are correctly identified in the generated summary.

The use of multiple ROUGE scores allows us to determine the model's effectiveness at different levels of n-grams and linguistic structures.

g) BLEU Score: The BLEU score is utilized to evaluate the precision of our model by comparing the generated summary against reference summaries. This metric measures the precision of n-grams in the generated summary compared to reference summaries. It provides a percentage indicating the proportion of correctly generated n-grams.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 13, May 2026
Impact Factor: 8.2

IV. RESULT AND DISCUSSION

A. Result

Our analyses focused on the application of extractive dialogue summarization using a integration of topic modeling and KeyBERT for keyword extraction, and advanced natural language processing techniques. Table I provides a detailed overview of the performance metrics obtained from our proposed summarization model. As explained in the section III, these metrics include ROUGE-1, ROUGE-2, ROUGE-L, and Bleu scores.

Methods	Rouge-1			Rouge-2			Rouge-L			Bleu
	R	P	F	R	P	F	R	F	R	
Approach 1	53.57	29.41	37.97	39.47	20.00	26.54	50.00	27.45	35.44	19.05
Approach 2	64.28	40.00	49.31	52.63	31.25	39.21	50.00	27.45	46.57	29.87

Table 1 Evaluating Our Model's Dialogue Summarization Performance with ROUGE and BLEU Scores

In Figure 5, we present the ROUGE scores for a particular dialogue sample evaluated under Approach 1. The figure includes detailed metrics such as ROUGE-1, ROUGE-2, and ROUGE-L, showcasing the recall, precision, and F1 scores.

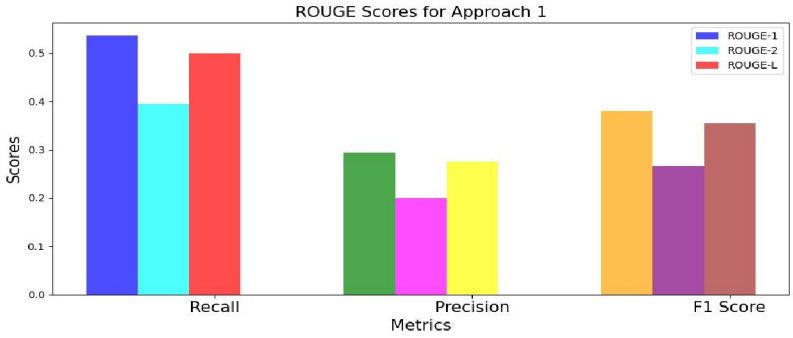

Figure 5 Rouge Scores for Approach 1 (Topic Modeling)

In Figure 6, we present the ROUGE scores for a particular dialogue sample evaluated under Approach 2.The figure includes detailed metrics such as ROUGE-1 (unigrams), ROUGE-2 (bigrams), and ROUGE-L (Longest Common Subsequence), showcasing the recall, precision, and F1 scores.

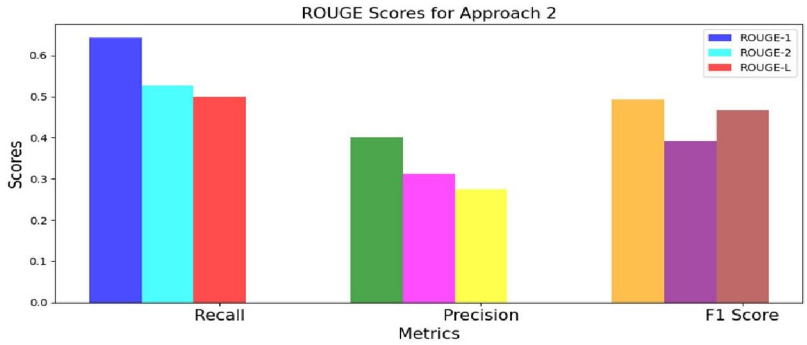

Figure 6 Rouge Scores for Approach 2 (KeyBERT)

The results and subsequent discussion highlight key findings:

1. Topic Modeling vs. KeyBERT:

Time Complexity: KeyBERT demonstrated faster processing times compared to topic modeling, making it more efficient for real-time applications.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 13, May 2026 Impact Factor: 8.2

Summary Quality: Both KeyBERT and topic modeling, particularly Latent Dirichlet Allocation (LDA), exhibit strong performance in generating quality summary sentences, showcasing their effectiveness in different dialogue scenarios.

2. Bert Embedding and Sentence Selection: Semantic Representation: Bert embeddings effectively captured the semantic nuances of sentences, contributing to improved similarity measurement.

PageRank Algorithm: PageRank-based sentence selection, considering both semantic similarity and keyword relevance, significantly enhanced summary coherence.

3. Integration of Speaker Roles:

Impact on Coherence: The absence of speaker role integration did not significantly affect coherence, suggesting that our approach effectively handled the conversational dynamics without explicitly using speaker roles.

4. Bonus Score with Keyword Matching: Enhanced Relevance: The incorporation of a bonus score for sentences containing extracted keywords resulted in summaries which were not only coherent but also more contextually relevant.

B. Comparative Analysis

The paper presents a comparative evaluation of our proposed approach against state-of-the-art models including BERT, GPT-2, and XLNet[31] in the task of dialogue summarization. We report ROUGE-1, ROUGE-2, ROUGE-L scores for each model to assess their summarization performance. Results indicate that while BERT, GPT-2, and XLNet exhibit relatively consistent performance across ROUGE metrics, our Approach 1 showcases competitive results, particularly in ROUGE-1 and ROUGE-2 scores. Specifically, Approach 1 achieves a ROUGE-1 score of 53.57% and a ROUGE-2 score of 29.41%. Approach 2 demonstrates further improvement, outperforming the baseline models across all ROUGE metrics, with notable gains in ROUGE-1, ROUGE-2, and ROUGE-L. Approach 2 achieves a ROUGE-1 score of 64.28%, marking a significant increase from BERT, GPT-2, and XLNet which achieved ROUGE-1 scores of 52.63%. Similarly, Approach 2 achieves a ROUGE-2 score of 40.00%, surpassing the baseline models' scores of 15.38%. In addition to ROUGE scores, we evaluate the effectiveness of our model and the baseline models using the BLEU score, which measures the similarity between the generated summary and the reference summary based on n-gram overlaps. Our Approach 2 achieves a BLEU score of 64.28%, indicating a significant improvement over BERT (52.63%), GPT2 (52.63%), and XLNet (52.63%). Compared to Approach 1 (39.47%), Approach 2 demonstrates a substantial increase in BLEU score by 24.81%. This significant improvement in BLEU score for Approach 2 highlights its superior ability to generate summaries that closely match the reference summaries compared to the baseline models and our previous approach. Therefore, our Approach 2 emerges as the most effective method for dialogue summarization tasks, offering better performance in terms of both ROUGE and BLEU scores.

Methods	Rouge-1			Rouge-2			Rouge-L			Bleu
	R	P	F	R	P	F	R	F	R	
BERT	64.28	15.38	24.82	52.63	10.05	16.87	60.71	14.52	23.44	8.43
GPT-2	64.28	15.38	24.82	52.63	10.05	16.87	60.71	14.52	23.44	8.43
XLNET	64.28	15.38	24.82	52.63	10.05	16.87	60.71	14.52	23.44	8.43
Approach 1	53.57	29.41	37.97	39.47	20.00	26.54	50.00	27.45	35.44	19.05
Approach 2	64.28	40.00	49.31	52.63	31.25	39.21	50.00	27.45	46.57	29.87

Table 2 Comparison of ROUGE and BLEU Scores: Our Model vs. BERT, GPT-2, XLNet

C. Discussion

• Efficiency vs. Quality Trade-off: The choice between KeyBERT and Topic Modeling involves a trade-off between efficiency and summary quality. Depending on the application requirements, one can opt for faster processing times or prioritize higher Rouge and Bleu scores.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 13, May 2026 Impact Factor: 8.2

• Dynamic Nature of Dialogues: Our approach effectively handled the dynamic and unstructured nature of dialogues. Speaker role information, while potentially beneficial, did not emerge as a critical factor, suggesting adaptability to various dialogue scenarios.

• Domain-Specific Adaptations: The flexibility of our approach allows for easy adaptation to specific domains. Finetuning the models on domain-specific datasets could further enhance the relevance of the generated summaries.

V. CONCLUSION

In conclusion, our refined approach for extractive dialogue summarization, integrating topic modeling and KeyBERT for keyword extraction, along with BERT embeddings and PageRank-based sentence selection, has demonstrated remarkable adaptability and efficiency. The flexibility to choose between KeyBERT and topic modeling depends on the nature and nuances of the given dialogue context. The incorporation of BERT embeddings ensures a nuanced understanding of sentence semantics, and the PageRank algorithm, complemented by keyword relevance, enriches sentence selection. Moreover, the bonus score addition for sentences containing extracted keywords enhances the informativeness of the final summary.

Our method, although extractive in nature, achieves a fine balance between performance and computational efficiency. The comparative advantage of each component in our pipeline allows users to tailor the summarization process to their specific needs. This adaptability, combined with competitive Rouge scores and bleu scores, positions our methodology as a promising solution for extractive dialogue summarization in diverse applications.

REFERENCES

[1]. P.k. Biswas, and A. Lakubovich "*Extractive Summarization of Call Transcripts*, " in IEEE Access, vol. 10,pp.119826-119840, Nov. 2022.

[2]. B. Ma, H. Sun, J. Wang, Q. Qi and J. Liao "*Extractive Dialogue Summarization Without Annotation Based on Distantly Supervised Machine Reading Comprehension in Customer Service*, " in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 30, pp.87–97, 2022.

[3]. H. Lin et al, "*Topic-Oriented Dialogue Summarization*, " in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 31, pp. 1797–1810, 2023.

[4]. Y. Liu and M. Lapata, "*Text summarization with pre-trained encoders*, " in Proc. EMNLP-IJCNLP, 2019, pp. 3730–3740.

[5]. J. Shin, S. -B. Park and H. -J. Song, "*Token-Level Fact Correction in Abstractive Summarization*, " in IEEE Access, vol. 11, pp. 1934-1943, 2023.

[6]. H. Zhu, L. Dong, F. Wei, B. Qin, and T. Liu, "*Transforming Wikipedia Into Augmented Data for Query-Focused Summarization*, " in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 30, pp. 2357-2367, 2022.

[7]. Alexandra N. Uma, Dmitry Sityaev, "*Comparing methods for extractive summarisation of call centre dialogue*, " 2022, arXiv:2209.02472.

[8]. Rubayyi Alghamdi, Khalid Alfalqi, "*A Survey of Topic Modeling in Text Mining*, " in (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 6, 2015.

[9]. Z. Wu, R. Koncel-Kedziorski, M. Ostendorf, and H. Hajishirzi, "*Extracting summary knowledge graphs from long documents*, ", 2020, arXiv:2009.09162.

[10]. M. Zhong, P. Liu, Y. Chen, D. Wang, X. Qiu, and X. Huang, "*Extractive summarization as text matching*, " in Proc. 58th Annu. Meeting Assoc. Comput. Linguistics, 2020, pp. 6197–6208.

[11]. V. Gupta and G. S. Lehal, X. Qiu, and X. Huang, "*A survey of text summarization extractive techniques*, " J. Emerg. Technol. Web Intell., vol. 2, no. 3, pp. 258–268, Aug. 2010.

[12]. S. Verma and V. Nidhi, "*Extractive summarization using deep learning*, " 2017, arXiv:1708.04439.

[13]. D. Miller, "*Leveraging BERT for extractive text summarization on lectures*, " 2019, arXiv:1906.04165.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 13, May 2026
Impact Factor: 8.2

[14]. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "*BERT: Pre-training of deep bidirectional transformers for language understanding*, " 2018, arXiv:1810.04805.
[15]. Y. Liu, "*Fine-tune BERT for extractive summarization*, " 2019, arXiv:1903.10318.
[16]. S. Narayan, S. B. Cohen, and M. Lapata, "*Ranking sentences for extractive summarization with reinforcement learning*, " inProc. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol., 2018, pp. 1747–1759.
[17]. Z. Wu, R. Koncel-Kedziorski, M. Ostendorf, and H. Hajishirzi, "*Extracting summary knowledge graphs from long documents*, " 2020, arXiv:2009.09162.
[18]. N. Franciscus, X. Ren, and B. Stantic, "*Dependency graph for short text extraction and summarization*, " J. Inf. Telecommun., vol. 3, no. 4, pp. 413–429, Apr. 2019.
[19]. Q. Zhou, N. Yang, F. Wei, S. Huang, M. Zhou, and T. Zhao, "*Neural document summarization by jointly learning to score and select sentences*, " in Proc. Assoc. Comput. Linguistics, 2018, pp. 654–663.
[20]. J. Cheng and M. Lapata, "*Neural summarization by extracting sentences and words*, " in Proc. Assoc. Comput. Linguistics, 2016.
[21]. K. Jezek and J. Steinberger, "*Automatic text summarization*, " in Proc. Znalosti. Princeton, NJ, USA: Citeseer, 2008, pp. 1–12.
[22]. J. Neto, A. A. Freitas, and C. A. Kaestner, "*Automatic text summarization using a machine learning approach*, " in Proc. Brazilian Symp. Artif. Intell.Berlin, Germany: Springer, 2002, pp. 205–215.
[23]. K. Kaikhah, "*Automatic text summarization with neural networks*, " in Proc. 2nd Int. IEEE Conf. Intell. Syst., Jun. 2004, pp. 40–44.
[24]. L. Suanmali, N. Salim, and M. S. Binwahlan, "*Fuzzy logic based method for improving text summarization*, " 2009, arXiv:0906.4690.
[25]. R. Nallapati, F. Zhai, and B. Zhou, "*Summarunner: A recurrent neural network based sequence model for extractive summarization of documents*, " in Proc. AAAI Conf. Artif. Intell., 2017, pp. 3075–3081.
[26]. J. Xu and G. Durrett, "*Neural extractive summarization with syntactic compression*, " in Proc. EMNLP-IJCNLP, 2019, pp. 3292–3303.
[27]. Q. Zhou, N. Yang, F. Wei, S. Huang, M. Zhou, and T. Zhao, "*A joint sentence scoring and selection framework for neural extractive document summarization*, " IEEE ACM Trans. Audio Speech Lang. Process., vol. 28, pp. 671–681, Jan. 2020. [Online].
[28]. X. Zhou et al., "*Multi-turn response selection for chatbots with deep attention matching network*, " in Proc. Assoc. Comput. Linguistics, 2018, pp. 1118–1127.
[29]. D. Wang, P. Liu, Y. Zheng, X. Qiu, and X. Huang, "*Heterogeneous graph neural networks for extractive document summarization*," in Proc. Assoc. Comput. Linguistics, 2020, pp. 6209–6219.
[30]. J. Xu, Z. Gan, Y. Cheng, and J. Liu, "*Discourse-aware neural extractive text summarization*, " in Proc. Assoc. Comput. Linguistics, 2020, pp. 5021–5031.
[31]. N. Mishra, M. Patel, P. Garg, and A. Varshney, "LLM aided semi-supervision for efficient Extractive Dialog Summarization," in Findings of the Association for Computational Linguistics: EMNLP 2023, pp. 10002–10009, Singapore, 2023.
[32]. F. Kirstein, T. Ruas, and B. Gipp, "CADS: A Systematic Literature Review on the Challenges of Abstractive Dialogue Summarization," Journal of Artificial Intelligence Research, vol. 82, pp. 1–60, 2025.
[33]. T. Le and A. T. Luu, "Extractive Summarization with Text Generator," in Proc. Annual Conf. of the North American Chapter of the Association for Computational Linguistics (NAACL), pp. 157–174, Mexico City, 2024.
[34]. Y. Wang, "Research on the TF-IDF Algorithm Combined with Semantics for Automatic Extraction of Keywords from Network News Texts," Journal of Intelligent Systems, vol. 33, no. 1, pp. 1–14, 2024.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 13, May 2026
Impact Factor: 8.2

[35]. R. Jain, P. Singh, and S. Puri, "Summarization of Daily News Using TextRank and TF-IDF Algorithm," in Proc. International Conference on Wireless and Satellite Systems (WiSSCoN 2023), Springer, pp. 312–323, 2024.

[36]. X. Liu and Y. Xu, "Learning to Rank Utterances for Query-Focused Meeting Summarization," in Findings of the Association for Computational Linguistics: ACL 2023, pp. 5765–5777, Toronto, Canada, 2023.

[37]. Y. Hu, T. Chen, F. Galley, and M. Zhu, "MeetingBank: A Benchmark Dataset for Meeting Summarization," in Proc. 61st Annual Meeting of the Association for Computational Linguistics (ACL 2023), pp. 13594–13611, Toronto, Canada, 2023.

[38]. H. Kim, M. Cho, and S. Na, "ExplainMeetSum: A Dataset for Explainable Meeting Summarization Aligned with Human Intent," in Proc. 61st Annual Meeting of the Association for Computational Linguistics (ACL 2023), pp. 8799–8815, Toronto, Canada, 2023.

[39]. S. Lamsiyah, A. El Mahdaouy, B. Espinasse, and S. E. Ouatik, "Unsupervised extractive multi-document summarization method based on transfer learning from BERT multi-task fine-tuning," Journal of Information Science, vol. 49, no. 1, pp. 164–182, 2023.

[40]. B. Bhargav, S. Kumari, R. Sharma, and P. Tiwari, "Graph-Based Extractive Text Summarization Sentence Scoring Scheme for Big Data Applications," Information, vol. 14, no. 9, p. 472, MDPI, 2023.