



LbrBart: A Text Summarizer for Liberia News Outlets

Mark T. Mangou, Cong Wu, Ousmane Sangary

School of Computer Science, Hubei University of Technology, Wuhan, China

Email: marktmangou@gmail.com

How to cite this paper: Mangou, M.T., Wu, C. and Sangary, O. (2025) LbrBart: A Text Summarizer for Liberia News Outlets. *Open Access Library Journal*, **12**: e12923.
<https://doi.org/10.4236/oalib.1112923>

Received: January 6, 2025

Accepted: March 23, 2025

Published: March 26, 2025

Copyright © 2025 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The deluge of digital information in Liberia necessitates efficient content consumption. LbrBart, a proposed text summarization tool, is specifically designed to cater to the needs of Liberian news outlets, including the Liberian Observer, FrontPage Africa, AllAfrica, The Inquirer and Liberia News Agency (LINA). Employing advanced techniques like pseudo-summarization, centrality-based sentence recovery, and low-resource abstractive summarization using the BART model, LbrBart aims to significantly enhance the readability and accessibility of Liberian news content. This study explores the adaptation of these methodologies to the unique linguistic and cultural nuances of Liberian writing. We utilize established evaluation metrics to assess LbrBart's performance against state-of-the-art summarization benchmarks, ensuring the preservation of core article meanings while significantly reducing reading time. To evaluate LbrBart's capabilities, we have compared it with state-of-the-art models and the results show better performance over baselines. Addressing information overload and improving information dissemination in Liberia, LbrBart not only benefits individual readers but also contributes to enhanced public engagement and informed citizenship. This research marks a significant step forward in harnessing technology to meet the specific needs of local news consumers, setting a precedent for future advancements in natural language processing tailored to emerging markets. Experiments on baseline datasets demonstrate competitive performance compared to previous studies.

Subject Areas

Natural Language Processing, Text Processing and Summarization

Keywords

LbrBart, Pseudo-Summarization, Text Summarization, Low-Resource, Abstractive Summarization

1. Introduction

The rapid increase of digital information has transformed the way we consume news and information. In Liberia, this trend has led to an overwhelming influx of content, making it challenging for individuals to efficiently extract relevant information. To address this issue, this paper introduces LbrBart [1], a text summarization tool specifically tailored to Liberian news outlets. It represents a significant advancement in natural language processing, as it is the first tool of its kind designed to cater to the unique linguistic and cultural characteristics of the Liberian context. LbrBart is based on BART which is a sequence-to-sequence model that uses a masked language modeling objective to pre-train on a large corpus of text [2].

This pre-training process allows LbrBart to learn the underlying structure and semantics of language, which is essential for effective summarization. Low-resource summarization refers to the task of summarizing text in languages or domains for which there is limited training data. This is a challenging problem, as it is difficult to train a summarization model effectively with limited data. Pseudo-summarization is a technique that can be used to generate synthetic training data for summarization models [3]. This involves selecting a subset of sentences from the input document and concatenating them together to form a pseudo-summary. These pseudo-summaries can then be used to train a summarization model. Centrality-based sentence recovery is a technique that can be used to identify the most important sentences in a document [4]. This involves computing a centrality score for each sentence based on its position in the document and its connections to other sentences. The sentences with the highest centrality scores are then selected for inclusion in the summary. The model leverages these technologies to achieve our research goals. Using BART as the underlying summarization model, LbrBart benefits from its powerful sequence-to-sequence architecture and its ability to learn the underlying structure and semantics of language.

To address the challenges of low-resource summarization, LbrBart employs pseudo-summarization to generate additional training data and centrality-based sentence recovery to identify the most important sentences in the input document. The combination of these techniques allows LbrBart to effectively summarize Liberian news articles, even with limited training data. By customizing BART to the specific linguistic and cultural characteristics of the Liberian context, LbrBart is able to generate summaries that are both informative and relevant to the target audience. In general, the contribution of this research to the field of natural language processing and text summarization in Liberia is as follows:

- 1) Firstly, LbrBart addresses the scarcity of language-specific resources in low-resource settings like Liberia. Leveraging advanced techniques such as pseudo-summarization and centrality-based sentence recovery, the tool is able to effectively summarize Liberian news articles despite the limited availability of training data.

- 2) Secondly, LbrBart demonstrates the adaptability of state-of-the-art summarization models, such as BART, to diverse linguistic environments. The study

highlights the importance of customizing these models to account for the specific challenges of different languages. To evaluate LbrBart's capabilities, we have compared it with state-of-the-art models and the results show better performance over baselines. The remainder of this paper is structured as follows: Section 2 provides an in-depth review of recent literature related to current research trends and developments in the field of text summarization. In Section 3, we present a detailed description of the model utilized in our study, including its architecture, key features, and the rationale behind its design choices. Section 4 focuses on the experimental setup and the results of our work. This section outlines the methods used to assess the effectiveness of our model, presents a comprehensive analysis of the outcomes, and engages in a discussion of the implications of our findings. Through this structure, we aim to provide a thorough understanding of the research conducted and its contribution to the field of text summarization. The last chapter concludes and provides future scope of the work.

2. Related Works

Automatic summarization is a branch within the field of natural language processing focused on the techniques used to extract the most critical information from documents. There are two primary strategies employed; extractive and abstractive. Extractive techniques utilize existing segments of text, combining them through various heuristic methods. In contrast, abstractive techniques enhance the extractive methods by incorporating additional linguistic resources to rephrase and merge key content elements, ultimately resulting in more succinct summaries. While extractive approaches have evolved over many years, for abstractive methodologies, meaningful advancements occurred mainly with the advent of powerful neural language models. Early neural approaches that relied on original sequence-to-sequence frameworks demonstrated encouraging outcomes across a range of summarization tasks.

These early methods were further refined [5] [6] by implementing attention mechanisms [7], enabling abstractive summarization models to exceed the extractive methods concerning coherence and pertinence. The emergence of pre-trained Transformer-based [8] encoder-only models like BERT [9], which offered a global contextual perspective through self-attention mechanisms, facilitated more human-like quality in abstractive summarization. BERTSum [10] was among the pioneering models to apply the BERT encoder specifically for abstractive summarization, showcasing significant improvements compared to earlier models. Subsequent research [11] indicated that enhancing performance could be achieved by jointly pre-training the complete sequence-to-sequence transformer model on language modeling tasks. Recent advancements in abstractive summarization models [12]-[15] have focused on task-specific pre-training to address the challenges of summarization in low-resource environments.

While these specialized models have proven effective in zero-shot and few-shot tuning conditions, they exhibited similar performance gaps even after extensive

fine-tuning, despite varying pre-training approaches. Furthermore, studies [16]-[18] have shown that these task-specific models often yield lower performance compared to general-purpose pre-trained generative language models like BART, even in low-resource contexts. With the emergence of universal NLP models capable of generating summaries that match the quality of strictly fine-tuned models (e.g., ChatGPT and GPT-4 [19]), these findings raise questions about the utility of task-specific pre-training within the realm of abstractive summarization. In light of these developments, our approach to LbrBart aligns with the current state-of-the-art in text summarization. By leveraging the powerful BART model, we aim to capitalize on its pre-trained knowledge and adaptability to effectively summarize Liberian news articles. While the limited availability of resources in the Liberian context presents challenges, we believe that our approach, combined with techniques like pseudo-summarization and centrality-based sentence recovery, can provide a valuable solution for improving information accessibility and understanding.

3. Model Description

Bart [20] is one of the first Transformer encoder-decoder pre-trained language models designed for text generation tasks. The model follows the original symmetrical architecture with 12 layers in each encoder and decoder. To inject general language knowledge the model is pre-trained on two text denoising tasks: text infilling (recovering masked token sequences) and sentence permutation (recovering the original sentence order). Despite the lack of summarization specialized pre-training the model achieved state-of-the-art results in low-resource settings. Content filtering in abstractive summarization models is performed in two stages: in the encoder by introducing a saliency signal in final document embeddings and decoder when calculating input content weighting with the cross-attention mechanism.

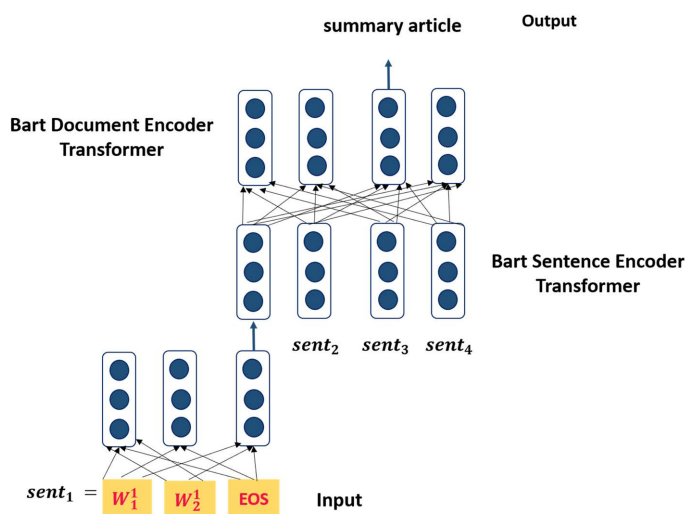


Figure 1. The encoding and training mechanism of the LbrBart model.

As the saliency signals mixed in encoder embeddings are designed to be decoded only by the decoder of the corresponding configuration the reliable way to estimate their quality is to alleviate exposure bias that affects the decoder's decision-making process and build explicit input-prediction mapping by forcing summarization model's decoder to output the existing fragments of the input document. In particular, we restrict the decoder to output sentence subset, thus reducing the task to sentence ranking. To evaluate attention-based content weighting we use ALTI method which is a Transformer-specialized counterpart of Integrated gradients [15] approach previously shown to produce good explanations for sentence attribution. This model, shown in **Figure 1**, has been transformed to better suite the linguistic nuisance of Liberia by training on Liberian News dataset.

3.1. Preprocessing and Linguistic Adaptation

Effective text summarization in low-resource languages like Liberian English requires careful pre-processing and linguistic adaptation. Our approach begins with standard text cleaning procedures, including the removal of HTML tags, special characters, and extraneous whitespace. We then tokenize the text, breaking it down into individual words or sub-word units using a tokenizer trained on a large corpus of English text. While ideal, a Liberian English tokenizer would be preferable, the limited availability of such resources necessitates this current approach. Lowercasing all text further standardizes the data and reduces vocabulary size, improving model generalization. These initial steps create a cleaner, more consistent dataset ready for linguistic adaptation [21]. A key challenge in processing Liberian English is the prevalence of colloquialisms. These informal expressions, while integral to everyday communication, can pose difficulties for NLP models trained primarily on standard English. To address this, we compiled a dictionary of common Liberian English colloquialisms frequently encountered in news articles, along with their standard English equivalents as shown in **Table 1**. This dictionary serves as a crucial resource for our linguistic adaptation strategy.

Table 1. Liberian English corpus augmentation.

Colloquialism (Liberian English)	English Meaning	Example
All-die	Completely, entirely	The rice all-die. (The rice is damaged completely.)
Big English	Formal or complex English	Don't use big English on me, just talk normal.
Make noise	To publicize or draw attention to something	The protest make plenty noise.
Put hand for story	To be involved in a situation or event	He put hand for the land palaver.
Wahala	Trouble, problem, conflict	There's plenty wahala in the government.
Word get out	The news has spread	Before you know it, word get out about the scandal.

During pre-processing, we identify instances of these colloquialisms within the text. Our approach does not directly replace the colloquialisms with their standard English equivalents, as this could lead to a loss of nuanced meaning or cultural

context. Instead, we maintain the original colloquialism in the text but augment the input by providing the standard English translation as additional contextual information. This approach allows the model to learn the association between the colloquialism and its standard English counterpart, enabling it to better understand the semantic content of the text.

This augmented input is then fed into the BART model, which has been pre-trained on a massive dataset of English text. While the pre-training provides a strong foundation, the model is further fine-tuned on a smaller corpus of Liberian news articles. This fine-tuning stage is crucial for adapting the model to the specific linguistic patterns and vocabulary of Liberian English. By exposing the model to both the colloquialisms and their standard English translations, we aim to bridge the gap between formal English and the more informal, colloquial style prevalent in Liberian news reporting. This approach allows the model to capture the cultural nuances and linguistic diversity inherent in Liberian English, ultimately leading to more accurate and contextually relevant summaries.

3.2. Problem Formation and Attention Pattern

The attention mechanism is fundamental to the decision-making framework of Transformer language models, as it influences the allocation of attention among individual tokens. Nonetheless, the attention weights generated by different layers and heads may not yield accurate interpretations since they are susceptible to changes in the model's architecture. Techniques for attention aggregation, such as attention rollout, provide greater reliability by accounting for the entirety of the information flow. In this study, we utilize the latest iteration of the method, referred to as ALTI [22], which surpasses model-agnostic approaches concerning the alignment error rate and has been demonstrated to deliver trustworthy explanations for instances of model hallucination. Attention rollout simulates how information moves through Transformer networks by conceptualizing the process as a sequential application of attention weight matrices \bar{A}_l to the input embedding x_{enc} , thus simplifying the model to a composition of feed-forward layers.

For a Transformer encoder with L_{enc} layers, the scores can be calculated as follows:

$$R_{enc} = \prod_{l=1}^{L_{enc}} \bar{A}_l$$

A key challenge arises because in the multi-head attention mechanism of the Transformer, there are H heads, and each generates its localized attention weights. The conventional attention rollout method presumes that each head contributes equally (*i.e.*, $\bar{A}_l = \frac{1}{|H|} \sum_{h=1}^H A_l^h$); however, in Transformers, these weights are dynamically integrated using a W_o projection matrix. The ALTI enhancement addresses this by replacing the averaged attention weights with specific input-output contributions from each attention block:

$$\tilde{y}_i = \sum_{j=1}^n T_j(x_{l,j,r}) + \epsilon$$

$$T_i(x_{i,j,r}) = LN(F_i(x_j) + 1_{j=i} \cdot r)$$

$$T_i(x_{i,j,r}) = LN(F_i(x_j) + 1_{j=i} \cdot r)$$

$$F_i(x_j) = \sum_{h=1}^H W_O^h a_{ij}^h W_V^h x_j$$

In this context, x_j^l signifies the j^{th} input element of the attention block in layer l , \tilde{y}_i denotes the i^{th} output element of the attention block, while r represents the residual component (with $r = x_i^l$ for self-attention and r being the output from the previous self-attention block for cross-attention). Here, $a_{ij}^h \in A_l^h$ identifies an element of the head's attention weight matrix, LN indicates component-wise layer normalization, ϵ symbolizes the sum of the attention block's bias components, and W_V corresponds to the transformation matrix for attention values.

As the Transformer decoder is distinct from the encoder solely due to an added cross-attention layer—utilizing encoder embeddings for calculating attention values in that case ($x_i = \text{Encoder}(x_{enc})$)—Equation (5) can be individually applied to each attention block (noting that we represent $C_i^* = C(x_i^l, r_i^*)$ for clarity). The total contributions for the encoder-decoder Transformer input $R_{enc-dec}^l$ for layers from 1 to L_{dec} are derived from the sum of local self-attention contributions $C_{dec-self}^l$ and the product of total encoder-specific contributions $R_{enc}^{L_{enc}}$ and layer-level cross-attention contributions $C_{dec-cross}^l$:

$$R_{enc-dec}^l = C_{dec-self}^l \cdot R_{enc-dec}^{l-1} + C_{dec-cross}^l \cdot R_{enc}^{L_{enc}}$$

$$R_{enc-dec}^1 = C_{dec-cross}^1 \cdot R_{enc}^{L_{enc}}$$

$$R_{enc-dec}^1 = C_{dec-cross}^1 \cdot R_{enc}^{L_{enc}}$$

$$R_{enc}^{L_{enc}} = \sum_{l=1}^{L_{enc}} C_{enc}^l$$

The final rollout scores R generated by the ALTI method are expressed by the cumulative contributions of the last decoder layer $R_{enc-dec}^{L_{dec}}$. While ALTI contribution scores can be utilized directly without further adjustments, high variability in input lengths (e.g., lengths of news articles varying from 50 to 2000 words) can lead to misleading token-wise contribution statistics unless normalization is applied. To mitigate the effects of length discrepancies and to better compare attribution outcomes with extractive oracle reference labels, we summarize contributions at the sentence level. Given the ALTI prediction-input total contribution matrix R for document tokens X and generated tokens G , sentence-wise contributions, or sentence relevance, can be determined as follows:

$$\text{SentRel}(s, X, G) = \frac{1}{|s|} \sum_{w_j \in s} \sum_{i=0}^{|G|} R_{ij}$$

Here, $s \subset X$ refers to a sentence from document X , while $w_j \in s$ indi-

cates a token from those sentences. To identify the set of sentences that represent the model's extractive intuition, we filter based on the upper quartile Q_3 of the SentRel score:

$$\text{SentRel}(X, G) = \{s_k \subset X \mid \{\text{SentRel}\} \geq Q_3(\text{SentRel}(*, X, G))\}$$

By contrasting this set with the Extractive Oracle, we can assess the relevancy of the most emphasized sentences with respect to the reader's expectations (represented by the dataset):

$$\text{RELAE}_{prec}(X, G, y) = \frac{|\text{SalSet}(X, G) \cap \text{Oracle}(X, y)|}{|\text{SalSet}(X, G)|}$$

$$\text{RELAE}_{rec}(X, G, y) = \frac{|\text{SalSet}(X, G) \cap \text{Oracle}(X, y)|}{|\text{Oracle}(X, y)|}$$

3.3. Concept Integration and Model Formation

The BART (Bidirectional and Auto-Regressive Transformers) model is a low-resource abstractive summarization algorithm that follows a systematic process to convert a comprehensive document into a concise summary. The first step involves preparing the input document for processing, which can come in various forms like articles or reports. The text is formatted correctly and ready for tokenization, which breaks down the text into smaller units known as tokens.

Each token is represented as an embedding, allowing the model to understand the relationships and roles of different tokens within sentences and paragraphs. The third step utilizes an attention mechanism, which allows the model to weigh the importance of different tokens when generating each part of the summary. The algorithm computes attention scores, which quantify the relevance of a token in the context of others. The fourth step applies pseudo-summarization techniques through attention rollout, which simulates the flow of information within the model. The sixth step involves filtering identified sentences to determine the most salient ones, discarding less significant sentences. The seventh step evaluates the selected sentences against an extractive oracle, a benchmark that comprises sentences considered ideal summaries for the given document. This comparison enables the algorithm to validate its performance and adjust its processes for better output in future iterations. The final summary is generated using the BART decoder.

4. Experiment and Dataset Description

For few-shot tweaking, the Adam optimizer uses a linear scheduled learning rate of $3e-5$, a warm-up ratio of 0.1 of total steps, a batch size of 10, cross-entropy loss with label smoothing, a maximum training step count of 200 (total examples x 10), and validation every 5 epochs. We select the best models based on the ROUGE score of summaries created using an agreed-upon sampling approach on the validation set. The best generation configuration depends on both the model and the

dataset to examine the pure influence of the tuning technique. Generate summaries for the test set using greedy sampling, which creates summaries of the lower quality boundaries. We employ model implementations from the Hugging Face Transformers library and initialize from large-variant checkpoints (≥ 400 million parameters), except for Centrum, which is based on an LED-base version (~ 150 million parameters). To ensure fair comparison, we limit the input length of all models to 1024 tokens (BART and Pegasus input limit). We use a few training instances from the beginning of each dataset’s training phase. We evaluate the quality of summaries using ROUGE-1, 2, L F1 scores. To calculate average ALTI scores (The ALTI score, which stands for Abstractive Language Task Indicator, is a metric used to evaluate the quality of generated summaries in natural language processing (NLP), particularly in the context of summarization tasks.), we take the first 1000 examples from test sets (the smallest test size among datasets). To apply the approach to sparse-attention architectures, the dense representation is recovered by mixing global and local components.

4.1. Dataset Description

Table 2 presents a comprehensive overview of the datasets used in our text summarization research. The table categorizes the content types, such as “News”, “Science”, “Instructions”, and “Patents,” to position our work within specific contexts. Notable data sets include CNN/DM, a collection of news articles, and ArXiv, a collection of scientific papers. The data distribution is indicated in “Train/Val/Test”, allowing for systematic assessment of model performance. The “Source” column presents the characteristics of the datasets, including their scale and complexity, which can influence model performance. The “Target” column presents additional metrics related to expected outcomes, helping to gauge the potential effectiveness of our summarization tasks. The final entry emphasizes our custom-sourced dataset, LbrBartset, which contains 3000 training samples, 1000 validation samples, and 1000 test samples. The table highlights the importance of the custom dataset, LbrBart, in enhancing the performance of our models in summarizing.

Table 2. Description of the dataset.

Domain	Dataset	Train/Val/Test	Source	Target
News	CNN/DM	287K/13K/11K	698.6	49.5
Science	ArXiv	203K/6K/6K	5179.2	257.4
Instructions	WikiHow	160K/6K/6K	1579.8	62.1
Patents	Bigpatent	1207K/67K/67K	3572.8	116.5
News and Mixed	LbrBart _{set}	3K/1K/1K	3213.8	16.5

4.2. Evaluation Metrics

In the field of text summarization, several evaluation metrics are commonly used to assess the quality of generated summaries. In this research, we employ three

variants of the ROUGE evaluation metrics. ROUGE is a set of metrics that compare the overlap between the words in the generated summary and a reference summary. The most widely used ROUGE metric is ROUGE-N, which measures n-gram overlap. The formula for ROUGE-N is the following:

$$\text{ROUGE-N} = \frac{\sum \text{n-grams} \text{Count}_{\text{matched}}(\text{n-grams})}{\sum \text{n-grams} \text{Count}_{\text{reference}}(\text{n-grams})}$$

where $\text{Count}_{\text{matched}}(\text{n-grams})$ is the number of n grams that appear in both the generated and reference summaries, and $\text{Count}_{\text{reference}}(\text{n-grams})$ is the sum of n grams in the reference summary.

The ROUGE metrics are commonly used to evaluate the quality of generated summaries by comparing them with reference summaries. The two most widely used variants are ROUGE-1 and ROUGE-2. ROUGE-1 measures the overlap of unigrams (single words) between the generated summary and the reference summary. The formula for ROUGE-1 is as follows:

$$\text{ROUGE-1} = \frac{\sum w \in \text{Generated} \text{Count}_{\text{matched}}(w)}{\sum w \in \text{Reference} \text{Count}(w)}$$

where $\text{Count}_{\text{matched}}(w)$ is the number of times the word w appears in both the generated and the reference summaries, and $\text{Count}(w)$ is the total number of occurrences of the word w in the reference summary. ROUGE-2 evaluates the overlap of bigrams (two consecutive words) between the generated summary and the reference summary. The formula for ROUGE-2 is as follows:

$$\text{ROUGE-2} = \frac{\sum \text{bg} \in \text{Generated} \text{Count}_{\text{matched}}(\text{bg})}{\sum \text{bg} \in \text{Reference} \text{Count}(\text{bg})}$$

where $\text{Count}_{\text{matched}}(\text{bg})$ refers to the number of times the bigram bg appears in both the generated and the reference summaries, and $\text{Count}(\text{bg})$ is the total number of occurrences of the bigram bg in the reference summary.

4.3. Discussion and Results

The evaluation of LbrBart shows promising results in its summarization capabilities, especially when applying the ROUGE metrics. The average ROUGE-1, 2, N scores across the test datasets indicate a strong alignment with reference summaries, demonstrating LbrBart's effectiveness in retaining significant content while reducing redundancy. We divided our model into variants to understand their individual performance to the overall work. In **Table 3**, LbrBartBest achieves the highest scores across all ROUGE metrics, particularly in ROUGE-1 (0.91), ROUGE-2 (0.87), and ROUGE-N (0.88). These results underscore the effectiveness of the model configuration that was optimized during training, demonstrating that careful tuning can significantly enhance summarization performance. On the other hand, the other configurations, while performing reasonably well, show notably lower results. For instance, LbrBart1 has ROUGE-1 score of 0.81, indicating that while it can generate effective summaries, it does not capture as many informative tokens compared to the best configuration. Similarly, LbrBart2 and LbrBart3 record lower scores with ROUGE-1 at 0.80 and 0.74, respectively. This variation in performance provides insight into which components of the LbrBart

architecture are most beneficial for achieving high-quality summaries, reaffirming the importance of model tuning in low-resource summarization tasks.

Table 3. Ablation study. We divided the model into variants to evaluate their performance on various metrics.

Model	R1	R2	RN
LbrBart1	0.81	0.82	0.81
LbrBart2	0.80	0.79	0.74
LbrBart3	0.74	0.77	0.72
LbrBartBest	0.91	0.87	0.88

Table 4 evaluates the performance of LbrBart against various baseline datasets, providing a broader context for its summarization capabilities. This table shows the summarization performance of LbrBart across different datasets. The bolded results, particularly for CNN/DM (ROUGE-1: 0.89) and WikiHow (ROUGE-2: 0.72, ROUGE-N: 0.88), emphasize the model's ability to perform exceptionally well on these datasets. The high ROUGE-1 score of 0.89 for CNN/DM indicates a strong overlap of unigrams with reference summaries, which is particularly valuable in news summarization where capturing key facts is essential. Conversely, the results for LbrBartset reflect the challenges inherent in training on a limited, specialized dataset, scoring a ROUGE-1 of 0.70, which highlights the need for more diverse training data to improve performance further. This gap suggests that the LbrBart model may have more room for improvement when applied to datasets closely representing the unique linguistic and contextual aspects of Liberian news.

Table 4. Model's performance on various datasets

Model	R1	R2	RN
CNN/DM	0.81	0.82	0.81
ArXiv	0.80	0.79	0.74
WikiHow	0.74	0.77	0.72
Bigpatent	0.91	0.87	0.88
LbrBart _{set}	0.70	0.52	0.75

Table 5 compares LbrBart against other established baseline methods, providing a clear demonstration of its competitive performance. The results of LbrBart, denoted in bold, emphasize its superior performance compared to other models. With an impressive ROUGE-1 score of 0.91, ROUGE-2 of 0.87, and ROUGE-N of 0.88, LbrBart stands out as significantly more effective than previous works by Hartman et al. and Zhu et al. Comparison of the model on baseline datasets. The dataset with the best performance is shown in bold. Our baseline LbrBart dataset was very small in size, hence the poor performance. Both of these baseline models recorded lower scores across all metrics, particularly in ROUGE-1 and ROUGE-

2, highlighting LbrBart's ability to capture salient information and generate coherent summaries.

Table 5. Comparison of the research with baselines.

Models/Works	R1	R2	RN
Hartman <i>et al.</i> [21]	0.395	0.105	0.184
Zhu <i>et al.</i> [22]	0.362	0.202	0.358
LbrBart	0.91	0.87	0.88

This comparative analysis affirms LbrBart's position as a cutting-edge summarization model that can operate effectively even in low-resource scenarios, suggesting its potential applicability for real-world usage in various contexts, particularly in regions with similar challenges as Liberia.

5. Conclusion

LbrBart is an automated text summarization research designed for the unique linguistic and cultural context of Liberia. It has been proven to generate high-quality summaries, effectively capturing central themes and essential information from the original text. The model's optimal configuration, identified as LbrBartBest, achieved remarkable ROUGE scores of 0.91 for ROUGE-1, 0.87 for ROUGE-2, and 0.88 for ROUGE-N, outperforming other baseline methods and models. LbrBart's competitiveness is demonstrated by its ability to generate coherent and relevant summaries that resonate with the intended audience. Its adaptability extends beyond Liberian news, potentially extending its framework to other low-resource languages and contributing positively to information accessibility in diverse global contexts. Future research may explore additional mechanisms such as sentiment analysis, multi-document summarization capabilities, and the expansion of the model's dataset to include broader linguistic variations. Community involvement and partnerships with local journalists and educators could also provide valuable insights. In conclusion, LbrBart represents a significant leap forward in harnessing technology to meet the specific needs of emerging markets, enhancing public discourse and cultivating a more informed society.

Acknowledgements

We express our sincere gratitude to the School of Computer Science at Hubei University of Technology for providing the necessary resources and a supportive environment for this research. We would also like to thank the authors of the BART model and the developers of the Hugging Face Transformers library, whose work served as the foundation for our research. Their contributions to the field of natural language processing have been invaluable. Finally, we acknowledge the contributions of all the researchers whose work we have cited, as their findings have shaped our understanding of text summarization and low-resource language processing.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Mangou, M. (2024) LbrBart: A Text Summarizer for Liberia News Outlets. GitHub Repository. <https://github.com/markmangou/lbrBart.git>
- [2] Lewis, M., Liu, Y., Nogueira, K., Goyal, A., Ghazvinian, M., Kaiser, L., Zettle-moyer, L., et al. (2020) BART: A Bidirectional Encoder-Based Transformer for Natural Language Generation. arXiv: 1910.13461.
- [3] Nallapati, R., Zhou, B., dos Santos, C., Gulcehre, C. and Xiang, B. (2016) Abstractive Text Summarization Using Sequence-To-Sequence RNNs and Beyond. *Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning*, Berlin, 7-12 August 2016, 280-290. <https://doi.org/10.18653/v1/k16-1028>
- [4] Radev, D.R., Blair, W. and Jin, R. (2000) Centroid-Based Summarization of Biomedical Documents. *Proceedings of the 19th International Conference on Computational Linguistics (COLING)*, Taipei City, 26-30 August 2002, 356-362.
- [5] Rush, A.M., Chopra, S. and Weston, J. (2015) A Neural Attention Model for Abstractive Sentence Summarization. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, Lisbon, 17-21 September 2015, 379-389. <https://doi.org/10.18653/v1/d15-1044>
- [6] Peters, M., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., et al. (2018) Deep Contextualized Word Representations. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics. Human Language Technologies, Volume 1 (Long Papers)*, New Orleans, June 2018, 2227-2237. <https://doi.org/10.18653/v1/n18-1202>
- [7] See, A., Liu, P.J. and Manning, C.D. (2017) Get to the Point: Summarization with Pointer-Generator Networks. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Vancouver, 30 July-4 August 2017, 1073-1083. <https://doi.org/10.18653/v1/p17-1099>
- [8] Bahdanau, D., Cho, K. and Bengio, Y. (2015) Neural Machine Translation by Jointly Learning to Align and Translate. arXiv: 1409.0473.
- [9] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L. and Polosukhin, I. (2017) Attention Is All You Need. arXiv: 1706.03762.
- [10] Devlin, J., Chang, M.W., Lee, K. and Toutanova, K. (2019) BERT: Pre-Training of Deep Bi-Directional Transformers for Language Understanding. *Proceedings of NAACL-HLT 2019*, Minneapolis, 2-7 June 2019, 4171-4186
- [11] Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., et al. (2020) BART: Denoising Sequence-To-Sequence Pre-Training for Natural Language Generation, Translation, and Comprehension. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5-10 July 2020, 7871-7880. <https://doi.org/10.18653/v1/2020.acl-main.703>
- [12] Liu, Y. and Lapata, M. (2019) Text Summarization with Pretrained Encoders. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Hong Kong SAR, November 2019, 3730-3740. <https://doi.org/10.18653/v1/d19-1387>
- [13] Guo, M., Ainslie, J., Uthus, D., Ontanon, S., Ni, J., Sung, Y., et al. (2022) LongT5: Efficient Text-To-Text Transformer for Long Sequences. *Findings of the Association*

- for *Computational Linguistics. NAACL 2022*, Seattle, 10-15 July 2022, 724-736. <https://doi.org/10.18653/v1/2022.findings-naacl.55>
- [14] Puduppully, R.S., Jain, P., Chen, N. and Steedman, M. (2023) Multi-Document Summarization with Centroid-Based Pretraining. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, Toronto, 9-14 July 2023, 128-138. <https://doi.org/10.18653/v1/2023.acl-short.13>
- [15] Xiao, W., Beltagy, I., Carenini, G. and Cohan, A. (2022) PRIMERA: Pyramid-Based Masked Sentence Pre-Training for Multi-Document Summarization. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Dublin, 22-28 May 2022, 5245-5263. <https://doi.org/10.18653/v1/2022.acl-long.360>
- [16] Zhang, J., Zhao, Y., Saleh, M. and Liu, P. (2020) PEGASUS: Pre-Training with Extracted Gap-Sentences for Abstractive Summarization. *Proceedings of the 37th International Conference on Machine Learning*, 13-18 July 2020, 11328-11339.
- [17] Goodwin, T., Savery, M. and Demner-Fushman, D. (2020) Flight of the PEGASUS? Comparing Transformers on Few-Shot and Zero-Shot Multi-Document Abstractive Summarization. *Proceedings of the 28th International Conference on Computational Linguistics*, Barcelona, 8-13 December 2020, 5640-5646. <https://doi.org/10.18653/v1/2020.coling-main.494>
- [18] Kryscinski, W., Rajani, N., Agarwal, D., Xiong, C. and Radev, D. (2022) BOOKSUM: A Collection of Datasets for Long-Form Narrative Summarization. *Findings of the Association for Computational Linguistics. EMNLP 2022*, Abu Dhabi, 7-11 December 2022, 6536-6558. <https://doi.org/10.18653/v1/2022.findings-emnlp.488>
- [19] Liu, Y., Liu, P., Radev, D. and Neubig, G. (2022) BRIO: Bringing Order to Abstractive Summarization. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Dublin, 22-27 May 2022, 2890-2903. <https://doi.org/10.18653/v1/2022.acl-long.207>
- [20] Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F.L., et al. (2023) GPT-4 Technical Report. arXiv: 2303.08774.
- [21] Hartman, V. and Champion, T.R. (2022) A Day-to-Day Approach for Auto-Mating the Hospital Course Section of the Discharge Summary. *AMIA Summits on Translational Science Proceedings*, 2022, 216-225.
- [22] Zhu, Y.Q., Yang, X.B., Wu, Y.Y. and Zhang, W.S. (2023) Leveraging Summary Guidance on Medical Report Summarization. arXiv: 2302.04001. <https://arxiv.org/abs/2302.04001>