computation

MDPI

*Article*

# Deep Learning-Based Short Text Summarization: An Integrated BERT and Transformer Encoder–Decoder Approach

**Fahd A. Ghanem [1,2,*], M. C. Padma [1], Hudhaifa M. Abdulwahab [3] and Ramez Alkhatib [4,*]**

[1] Department of Computer Science & Engineering, PES College of Engineering (Affiliated to University of Mysore), Mandya 571401, India; padmapes@gmail.com
[2] Department of Computer Science, College of Education-Zabid, Hodeidah University, Hodeidah P.O. Box 3114, Yemen
[3] Department of Computer Application, Ramaiah Institute of Technology (Affiliated to VTU), Bengaluru 560054, India; hudhaifa.alhimyari@gmail.com
[4] BMB Nord, Research Center Borstel, Parkallee 35, 23845 Borstel, Germany
* Correspondence: fahd.a.ghanem@gmail.com (F.A.G.); ralkhatib@fz-borstel.de (R.A.)

**Abstract:** The field of text summarization has evolved from basic extractive methods that identify key sentences to sophisticated abstractive techniques that generate contextually meaningful summaries. In today's digital landscape, where an immense volume of textual data is produced every day, the need for concise and coherent summaries is more crucial than ever. However, summarizing short texts, particularly from platforms like Twitter, presents unique challenges due to character constraints, informal language, and noise from elements such as hashtags, mentions, and URLs. To overcome these challenges, this paper introduces a deep learning framework for automated short text summarization on Twitter. The proposed approach combines bidirectional encoder representations from transformers (BERT) with a transformer-based encoder–decoder architecture (TEDA), incorporating an attention mechanism to improve contextual understanding. Additionally, long short-term memory (LSTM) networks are integrated within BERT to effectively capture long-range dependencies in tweets and their summaries. This hybrid model ensures that generated summaries remain informative, concise, and contextually relevant while minimizing redundancy. The performance of the proposed framework was assessed using three benchmark Twitter datasets—Hagupit, SHShoot, and Hyderabad Blast—with ROUGE scores serving as the evaluation metric. Experimental results demonstrate that the model surpasses existing approaches in accurately capturing key information from tweets. These findings underscore the framework's effectiveness in automated short text summarization, offering a robust solution for efficiently processing and summarizing large-scale social media content.

**Keywords:** attention mechanism; BERT; deep learning; short text summarization; transformer-based encoder–decoder; Twitter summarization

check for updates

Academic Editor: Zhiyong (Johnny) Zhang

Received: 3 February 2025
Revised: 4 April 2025
Accepted: 10 April 2025
Published: 12 April 2025

**Citation:** Ghanem, F.A.; Padma, M.C.; Abdulwahab, H.M.; Alkhatib, R. Deep Learning-Based Short Text Summarization: An Integrated BERT and Transformer Encoder–Decoder Approach. *Computation* **2025**, *13*, 96. https://doi.org/10.3390/computation13040096

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

The explosive growth of big data and widespread internet access has inundated people with vast amounts of online text. The ability to distill this deluge into digestible, informative summaries is not just a convenience; it is a necessity. As a result, there is a growing need for technological methods that can autonomously summarize text, providing users with quick access to important information without sacrificing the original content's meaning. However, manually summarizing large amounts of text is both costly and time-consuming. In this context, automatic text summarization (ATS) has emerged as a pivotal field of study, aiming to preserve the essence of the main text while eliminating unnecessary material to

provide concise summaries [1]. The development of robust summarization techniques has become a focal point in natural language processing (NLP) research [2].

ATS has been a foundational aspect of NLP research since its inception in the mid-20th century [3]. Over time, multiple ATS techniques have emerged, including extractive and abstractive strategies for both single- and multi-document summarization. Extractive methods involve selecting key sentences directly from the source text based on their grammatical and statistical features to form a summary; several examples of extractive techniques can be found in [4]. In contrast, abstractive methods reinterpret and rephrase information to convey the main points. Abstractive summarization is more challenging, as it demands a higher level of understanding to ensure that the summary is fluent, salient, coherent, accurate, and novel [5].

Recent advancements in DL have significantly transformed the landscape of text summarization [6], enabling the development of models that can understand and generate human-like summaries. Among these advancements, the transformer architecture [7] has emerged as a powerful framework due to its attention mechanisms, which allow models to focus on relevant portions of the input text. Additionally, the introduction of BERT [8] has emerged as a powerful tool for capturing context and semantics within text, leading to substantial improvements across various NLP tasks, including summarization. BERT's ability to understand the nuances of language enhances the quality of generated summaries by providing richer representations of the input text.

The literature reveals an expanding body of research focused on leveraging deep learning techniques for text summarization [9]. Notable studies have highlighted BERT's remarkable effectiveness for extractive summarization tasks, showcasing its ability to capture essential information from texts [10]. Additionally, the potential of transformer-based models for generating coherent summaries has been noted [11]. The integration of BERT with transformer-based models has been explored in various studies, revealing an enhanced ability to capture the nuances of language and generate coherent summaries [12].

Despite significant advancements, there remains a gap in research focused on integrating these two powerful frameworks for short text summarization. Current models often struggle to balance the trade-off between informativeness and conciseness, leading to suboptimal results. They frequently fail to capture deeper meanings and critical elements, resulting in fragmented or incomplete summaries. Moreover, while extensive research has been conducted on summarizing longer documents, short text summarization presents unique challenges, such as limited context and the need for more precise language models, making it a promising yet underexplored area in NLP.

This research aims to bridge this gap by proposing an integrated system architecture based on BERT and the TEDA enhanced with an attention mechanism for the purpose of extractive summarization of short Twitter text. The proposed method offers various advantages; due to the bidirectional nature of encoding, BERT captures the meaning of the entire tweet, allowing for better summary generation. The construction of the TEDA architecture enables the use of context information in both directions, ensuring coherent summary generation. By concentrating on important passages within the text, the attention mechanism further enhances the extractive summarization process. Additionally, long short-term memory (LSTM) networks are incorporated within BERT to capture long-range dependencies in tweets and their summaries. This integrated approach is specifically designed to accommodate the demands of the concise and casual language of Twitter, enabling scalability for large datasets and achieving innovative performance in extractive summarization. The main contributions of the framework can be summarized as follows:

- Developing a robust data cleaning process to remove non-essential elements from tweets, ensuring that the essential that the dataset is free of noise and irrelevant content.
- Implementation of preprocessing techniques, including tokenization, stop word removal, stemming, and lemmatization, specifically designed to optimize input data for short text summarization tasks.
- Proposing a novel framework integrating BERT with a TEDA enhanced with an attention mechanism to effectively summarize short-text content on Twitter.
- Evaluating the framework using key metrics, such as ROUGE scores, across three disaster tweet datasets such as Hagupit, SHShoot, and Hyderabad Blast to assess understanding, coherence, and informativeness.

The following is a summary of the research: Section 2 examines earlier research that used a variety of text-summarizing strategies. Section 3 formulates the research problem and provides a detailed description of the proposed framework, including an approach that integrates attention mechanisms with BERT and a transformer-based encoder–decoder structure. The analytical outcomes of the framework's performances are shown in Section 4. In Section 5, the research is concluded and the findings are discussed.

## 2. Related Works

In an era where information overload is a daily challenge, the ability to distill large volumes of text into concise summaries has become increasingly vital. ATS has become a significant task in the field of NLP, aimed at producing concise and coherent summaries of the key information contained in larger texts [13]. This section highlights key contributions in the field, particularly those relevant to our proposed framework. Recent studies [1–3,14] provide comprehensive reviews of various summarization techniques. Over the years, ATS techniques have evolved and been applied across various domains, such as summarizing news articles [6,15], legal document summarization [16], medical document summarization [17], and social media posts, including tweet summarization [18]. In particular, social media platforms, with their informal and abbreviated nature, have posed unique challenges for ATS, as explored by [1]. Their work highlights the need for more effective short text summarization on platforms like Twitter.

The rapid spread of information on social media platforms, particularly Twitter, which has become a major source for sharing news, updates, opinions, and real-time information on ongoing events, presents challenges in identifying relevant insights amid the overwhelming volume of tweets [19]. Several existing tweet summarization methods can be classified based on the events they focus on, such as political events summarization [20], natural disasters summarization [21–23], and sports events [24,25].

Recent studies [26–28] have explored the summarization methods of disaster event tweets, which contain situational insights and time-critical information shared by eyewitnesses and affected individuals. The approaches specific to disaster tweet summarization can be classified into abstractive methods [29–31] and extractive methods [22,26,27,32,33]. The paper by K. Rudra et al. [34] presented a hybrid extractive–abstractive approach for summarizing situational tweets during crises, aiming to provide emergency management with concise and relevant insights from social media data. This framework first constructs a graph with disaster-specific keywords as nodes and bigram relationships as edges, capturing key information relationships. It then selects tweets that maximize information coverage within this graph, enhancing summary quality and relevance for decision-makers in disaster response. Moreover, Rudra et al. [35] proposed a classification–summarization framework to extract critical situational information from microblog streams shared during disaster events. Ghanem et al. [1] pointed out the dynamic and fragmented nature of

tweets, which makes the summarization tasks difficult. However, among the various tweet summarization tasks, summarizing disaster-related tweets poses distinct challenges. These tweets are often unstructured, time-sensitive, and noisy, complicating the extraction of relevant information [1]. Given these challenges posed by the immense volume of data on Twitter, advanced summarization techniques are essential to extracting meaningful content from the overwhelming flood of posts.

The advent of DL has paved the way for the development of more sophisticated models for tweet summarization. One prominent model is BERT, introduced by Devlin et al. [8]. BERT revolutionized NLP by pre-training deep bidirectional representations that capture contextual word meanings from both directions in a text. This innovation has led to significant improvements in both extractive and abstractive summarization tasks. Liu and Lapata [10], who applied pre-trained encoders for text summarization, demonstrated how pre-trained encoders like BERT can identify important text segments that best represent the overall content, achieving state-of-the-art results.

BERT has proven effective for tweet summarization, especially in disaster scenarios. Building on BERT's capabilities, various models have been developed. For example, Garg et al. [33] proposed IKDSumm, an approach for automatic extractive disaster tweet summarization. This approach incorporates domain-specific knowledge and key phrases into a BERT-based model to produce high-quality summaries. Garg et al. [23] also introduced OntoDSumm, a framework that uses ontology-based feedback mechanisms to enhance the quality of tweet summaries. Additionally, ATSumm [29], another model, emphasizes the use of auxiliary information to improve summarization performance in crisis situations. Other approaches, such as BERTSum [36] and its variants, have demonstrated significant improvements in extractive summarization by capturing sentence-level representations and selecting the most relevant sentences from the source text. Furthermore, La Quatra and Cagliero [37] developed BART-IT, a sequence-to-sequence model specifically designed for summarizing Italian text.

Recent studies have introduced CNsum [38], a sequence-to-sequence model that employs BERT as an encoder and a transformer-based GPT-2 decoder to generate Chinese news headlines. This model addresses the complexities of non-English summarization by leveraging BERT's bidirectional contextual representations and GPT-2's generative capabilities. It operates through a two-step process: first, encoding input text into contextual embeddings, followed by decoding them into coherent summaries. Performance evaluations on the NLPCC2017 dataset highlight its effectiveness, demonstrating the adaptability of BERT-transformer architectures across various languages and domains. This research reinforces the significance of encoder–decoder frameworks while expanding their applicability to low-resource language settings.

Divya et al. [39] introduced a comprehensive framework that combines BERT-based extractive techniques with transformer-based abstractive summarization. Their method leverages SBERT (Sentence-BERT) to generate sentence embeddings for key phrase identification, followed by a transformer decoder to produce coherent summaries. By integrating extraction for source fidelity and abstraction for contextual flow, this approach strikes a balance between accuracy and fluency. The framework achieves state-of-the-art performance on datasets from news and legal domains, highlighting the effectiveness of merging BERT's contextual representations with transformer decoders. Their findings align with the principles of multi-stage summarization, reinforcing the value of an encoder–decoder architecture.

Recent research by Abadi and Ghasemian [40] has enhanced Persian text summarization by applying a three-phase fine-tuning and reinforcement learning (RL) method using the mT5 transformer model. To overcome challenges in low-resource environments, their approach fine-tunes mT5-base on Persian news articles through three stages: (1) domain adaptation to align with Persian syntax and semantics, (2) task-specific tuning to optimize for summarization, and (3) RL refinement with proximal policy optimization (PPO) to improve contextual accuracy and coherence. This combined approach surpasses previous Persian-focused models, achieving state-of-the-art ROUGE scores (ROUGE-1: 53.17, ROUGE-L: 44.13).

Another prominent model, the transformer-based encoder–decoder model, has also had a tremendous impact on summarization. Models of this nature, such as Vaswani et al.'s [7], employ a process where they use an encoder to encode the input sequence and a decoder to generate the summary. The mechanism of self-attention allows the model to impose weights on different elements of the input text so that it can output the summary more coherently. Current research has also started to look at how BERT can be combined with transformer-based encoder–decoder models. Zhang et al. [12] introduced a model that merges BERT with an attention mechanism to enhance the quality of generated summaries. Their results showed significant improvement in coherence as well as informativeness over classical approaches. To effectively extract important information from user comments during significant occurrences, Papagiannopoulou and Angeli [41] attempt to construct a social media text summarizing system. The suggested system makes use of transformer-based models, particularly the pre-trained T5 model, abstractive summarization, and transfer learning. The quality and relevancy of summaries are much improved by transformer-based approaches, according to the findings.

Kherwa et al. [42] suggested an approach that created contextual, semantically rich summaries for big text collections by combining BERTsum for extractive summarization with Longformer2Roberta for abstractive summarization. On three datasets—CNN, WikiSum, and Gigaword—the model demonstrated a notable improvement in performance over several levels of Rouge score and word mover distance.

Toprak and Turan [43] proposed a hybrid method that combined phrase grouping and transformer to perform automated document summarization. After being filtered for logical and spelling mistakes, the BBC News dataset was used to train the transformer model. After that, documents were broken down into sentences and the relationship was determined using the Simhash text similarity technique to test the model.

Murugaraj et al. [44] presented HistBERTSum-Abs, a unique approach to abstractive historical single-document summarizing. The technique outperformed deep learning-based techniques and LLMs in zero-shot factors by utilizing HistBERT, a unidirectional language model, and an encoder–decoder architecture.

Despite BERT's effectiveness in handling longer documents, its application to short text summarization on social media platforms like Twitter has not been as thoroughly studied. As a result, there is still a need for integrated approaches that leverage the representational power of BERT with the generative capabilities of transformer architectures. Existing research often overlooks the potential of combining these frameworks, which can lead to better handling of the unique challenges posed by Twitter short text summarization.

This study aims to address this gap by proposing an ATS framework that integrates BERT with transformer-based architectures, enhanced by LSTM networks for Twitter short text summarization, especially in disaster-relevant tweets. The framework aims to generate high-quality summaries that accurately reflect the main ideas of the source text. The significance of this research lies in its potential to enhance automated short summarization.

Table 1 provides a summary of the related works, highlighting the technologies used, results, advantages, and limitations for automatically summarizing brief texts. It offers a comprehensive overview of various approaches in this field.

**Table 1.** An overview of the relevant literature on automatic short text summarization.

| Study | Technology Used | Result | Advantages | Limitations |
| --- | --- | --- | --- | --- |
| Zhang et al. [45] | SummIt with large language models (LLMs). | Improved controllability and faithfulness of summaries with human-like writing by using feedback mechanisms. | Self-evaluation improves the quality and human-likeness of summaries. | Requires extensive human assessment and training to refine models. |
| Paul et al. [26] | GAN-based clustering for disaster data summarization. | AI-based clustering of tweets during disasters, demonstrating GAN's efficacy on large tweet datasets. | Effectiveness in handling large datasets, especially in crisis situations. | GAN-based method focused on disaster tweets, limiting broader applicability. |
| Garg et al. [29] | Auxiliary pointer generator network (AuxPGN) with key-phrase attention for ATSumm. | Achieved 4–80% improvement in ROUGE-N F1 scores across 13 disaster datasets compared to 10 state-of-the-art approaches. | Effectively addresses data sparsity using key phrases as auxiliary information; superior performance in disaster summarization. | Requires specialized key-phrase extraction process; potentially limited to disaster-related content. |
| Garg et al. [33] | IKDSumm framework for disaster-specific tweet summarization. | Demonstrated an approximate F1 improvement over current techniques across 12 disaster datasets. | Ensures coverage, relevance, and diversity in disaster-related tweet summaries. | Focused specifically on disaster-related tweets, potentially limited in general tweet summarization scenarios. |
| Garg et al. [23] | Ontology-based disaster tweet summarization. | Outperformed six state-of-the-art methods by 2–66% in terms of ROUGE-1 F1 score across 12 disaster datasets | Utilized domain knowledge via ontology for improved tweet categorization and summarization | Requires pre-defined domain-specific ontology; less adaptable to unseen or evolving categories |

## 3. Proposed Framework for Short Text Summarization

Automatic text summarization (ATS) is a complex challenge, particularly when applied to short, noisy, and contextually diverse texts such as tweets. Current approaches often struggle to strike a balance between compression and retention, leading to either overly generalized or incomplete summaries. The proposed framework, shown in Figure 1, integrates the strengths of bidirectional encoder representations from transformers (BERT) and transformer encoder–decoder architectures (TEDA) with an advanced attention mechanism and LSTM units. This design addresses the unique challenges of Twitter-based summarization, such as limited character counts, informal writing styles, and the need for high computational efficiency.

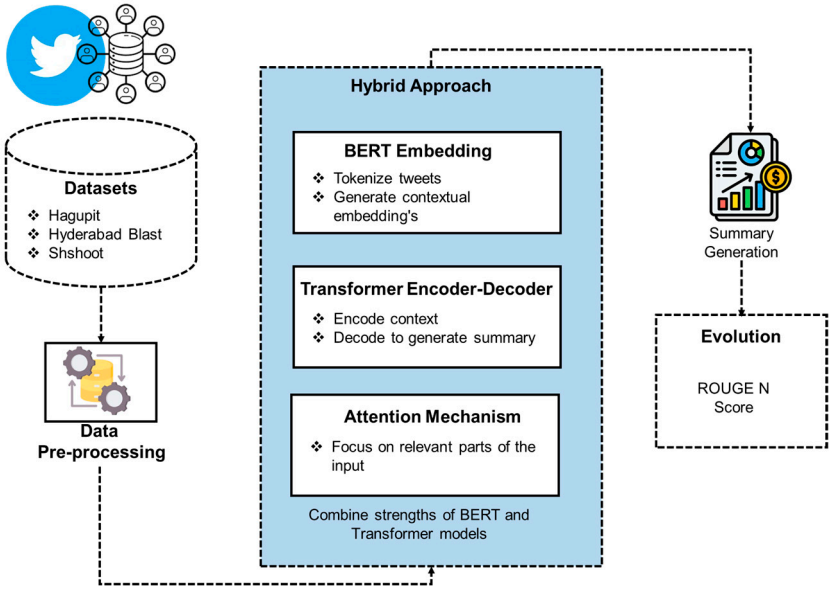



**Figure 1.** Proposed framework.

### *3.1. Data Collection and Preprocessing*

The initial phase of the framework focuses on collecting and preparing the dataset for summarization. Twitter's API is used to collect large volumes of tweets relevant to specific topics, such as disaster events or trending news. In this study, we used three disaster event datasets collected by [35] for experimentation.

- Typhoon Hagupit: A strong cyclone, namely, Typhoon Hagupit, hit the Philippines.
  - https://cse.iitkgp.ac.in/~krudra/typhoon_hagupit_dataset.html (accessed on 8 November 2024)
- Sandy Hook Shooting (SHShoot): An assailant killed 20 children and 6 adults at the Sandy Hook Elementary School in Connecticut, USA.
  - https://cse.iitkgp.ac.in/~krudra/sandyhook_shoot_dataset.html (accessed on 8 November 2024)
- Hyderabad Blast: Two bomb blasts in Hyderabad city of India.
  - https://cse.iitkgp.ac.in/~krudra/hyderabad_blast_dataset.html (accessed on 8 November 2024)

The collected raw data undergoes a preprocessing pipeline to ensure uniformity and minimize noise [46]. The preprocessing steps, as shown in Figure 2, are outlined below:

- Data Cleaning:
  Unnecessary components such as URLs, hashtags, mentions, and special characters are removed, and the text is converted to lowercase for consistency. Emojis and emoticons are either replaced or removed depending on their relevance.
- Tokenization and Subword Representation:
  BERT's tokenization process breaks down out-of-vocabulary terms into smaller sub-units, while special tokens mark the beginning and end of sequences.
- Handling Stop Words and Lemmatization:
  Common stop words are removed, and lemmatization is used to reduce words to their base forms, ensuring the preservation of meaning.
- Padding and Truncation:
  Tweets are padded to meet the input size requirements of BERT, and longer tweets are truncated to ensure they fit within the model's 512-token limit.

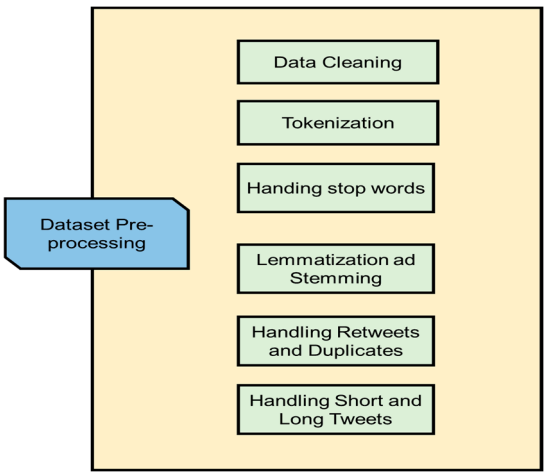



**Figure 2.** Tweet dataset pre-processing steps.

### *3.2. Hybrid BERT and Transformer Encoder–Decoder Architecture*

The backbone of the proposed framework is the integration of BERT and a transformer-based encoder–decoder architecture, as shown in Figure 3, enhanced with attention mechanisms and LSTM units for sequential dependency modeling.

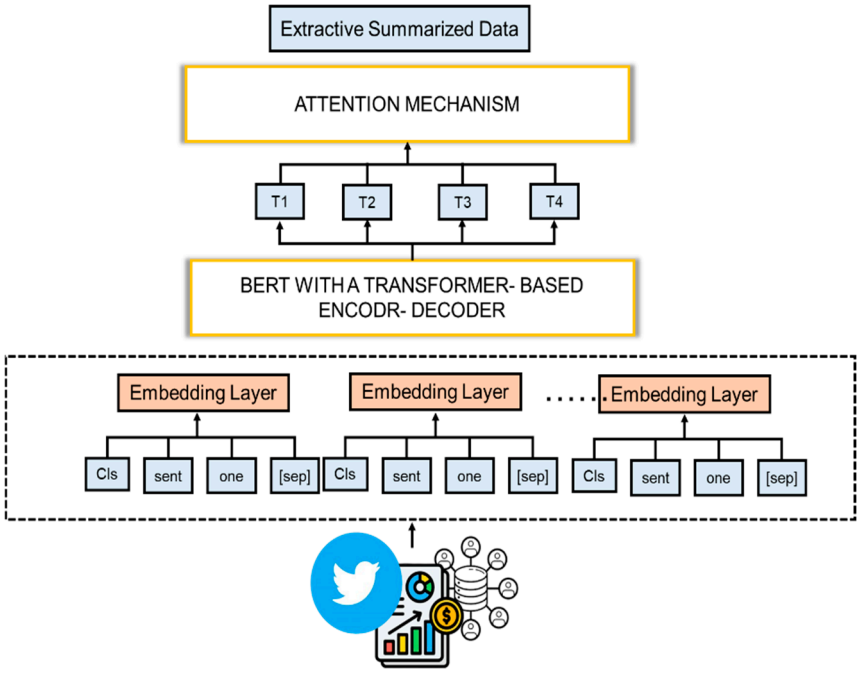



**Figure 3.** Hybrid BERT with a transformer-based encoder–decoder and attention mechanism.

#### 3.2.1. Contextual Embedding with BERT

- BERT's tokenizer processes the preprocessed tweets, generating token embeddings that capture the syntactic and semantic relationships within the text. The embeddings are context-aware and represent each token within the broader sentence-level context, ensuring that nuances specific to tweets are preserved [47].
- The embedding for each token, $E$, is generated by combining the token embedding $E_t$, segment embedding $E_s$, and positional embedding $E_p$, as represented by Equation (1):

$$E = E_t + E_s + E_{\mathrm{p}} \tag{1}$$

3.2.2. Transformer Encoder–Decoder with Attention

- The transformer encoder processes the embeddings produced by BERT, leveraging its self-attention mechanism to capture long-range dependencies and relationships between tokens [48]. The self-attention mechanism computes the importance of each token pair using Equation (2):

$$Attention(Q, K, V) = SoftMax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

Here, $Q$, $K$, and $V$ are the query, key, and value matrices, and $d_k$ is the dimension of the keys [49].

- The encoder's output is passed to the decoder, which generates summaries by processing the encoded context and iteratively predicting the next tokens in the sequence.

3.2.3. LSTM Integration for Sequential Learning

- LSTM units are incorporated into the architecture to enhance the model's ability to capture sequential patterns and temporal dependencies. This is especially important for tweets, which often contain time-sensitive information and rely on the sequence of words to convey meaning.

3.2.4. Attention Mechanism for Relevance Focus

- An integrated attention mechanism selectively focuses on the most relevant parts of the input text, enabling the model to prioritize essential information while filtering out redundant or irrelevant details. This ensures that the generated summaries are concise yet highly informative.

### *3.3. Summary Generation and Evaluation*

The final stage of the framework involves generating summaries and evaluating their quality. The decoder produces summaries in a logical and coherent structure, ensuring readability and retention of critical information. The generated summaries are evaluated using ROUGE scores, which measure the overlap of recall, precision, and F1 score between the generated summaries and reference texts. Specifically,

- ROUGE-1: Measures the overlap of unigrams.
- ROUGE-2: Measures the overlap of bigrams.
- ROUGE-L: Measures the longest common subsequences, reflecting fluency and coherence.

### *3.4. Advantages of the Proposed Framework*

The proposed hybrid framework offers several advantages:

- Contextual Understanding: BERT's bidirectional transformer effectively captures the nuanced semantics and syntactic structures in tweets, overcoming challenges posed by informal and noisy language.
- Relevance and Focus: The attention mechanism ensures that the summaries prioritize essential information, minimizing the inclusion of irrelevant content.
- Sequential Dependency Modeling: The integration of LSTM units enhances the model's ability to understand temporal relationships, critical for tweets with event-driven content.
- Adaptability to Twitter Data: The preprocessing pipeline and hybrid architecture are specifically tailored to handle the brevity and diversity of Twitter data, ensuring high-quality summaries that are both concise and informative.

### *3.5. Pseudocode*

The summarization process is outlined in Scheme 1.

```
BEGIN
Step 1: Preprocess tweets
    FOR each tweet in collected_tweets:
        REMOVE noise (hashtags, URLs, special characters)
        TOKENIZE tweet using BERT tokenizer
        LEMMATIZE and NORMALIZE tokens
        PAD or TRUNCATE tokenized tweet to fit model input size
Step 2: Generate embeddings using BERT
    FOR each preprocessed tweet:
        COMPUTE token embeddings, segment embeddings, and positional embeddings
        PASS embeddings through BERT layers to obtain contextualized representations
Step 3: Transformer encoder–decoder with LSTM
    INITIALIZE Transformer Encoder
    FOR each tweet embedding:
        APPLY self-attention in Encoder to capture dependencies
    INITIALIZE LSTM layer
    FOR each encoder output:
        PASS through LSTM to enhance sequential learning
        INITIALIZE Transformer Decoder
    FOR each encoded representation:
        GENERATE concise summary using decoder
# Step 4: Apply attention mechanism
    FOR each decoder step:
        COMPUTE attention scores over encoder outputs
        WEIGHT encoder outputs based on relevance
        SELECT key components for summary generation
# Step 5: Evaluate summaries
    FOR each generated summary:
        COMPUTE ROUGE-1, ROUGE-2, and ROUGE-L scores
        COMPARE with reference summaries for evaluation
OUTPUT extractive summaries capturing core tweet information
END
```

**Scheme 1.** Pseudocode: BERT and transformer encoder-decoder with attention.

The proposed framework effectively addresses the challenges of summarizing short texts on platforms like Twitter by combining the strengths of BERT and TEDA with advanced attention and sequential modeling capabilities.

## 4. Result and Discussion

In this section, we present the experimental results of our study on automatic short text summarization using the Hybrid BERT-LSTM model. We begin by outlining the experimental setup (Table 2), followed by a discussion of the performance metrics and baseline methods. The effectiveness of our model is evaluated through ROUGE scores across various metrics. We also compare our model with established summarization techniques and conclude with an interpretation of the implications of our findings.

**Table 2.** Experimental setup.

| Experimental Setup | Details |
|---|---|
| Model | Hybrid BERT-LSTM with transformer-based encoder–decoder architecture (TEDA) |
| Task | Automatic short text summarization on Twitter |
| Dataset | Hagupit, SHShoot, Hyderabad Blast tweet datasets |
| Hardware | Laptop running Windows 11 |
| Processor | Intel Core i7 Evo 13th Gen |
| RAM | 16 GB |
| Software Environment | Python 3.10.1 |
| Evaluation Metrics | Precision, recall, F1 score, ROUGE score [50] |
| Implementation code | https://github.com/NLPSTS/BERT-and-TEDA- (accessed on 4 April 2024) |

### *4.1. Baseline Methods*

In this study, we compare our proposed Hybrid BERT-LSTM framework with several established baseline methods, specifically for disaster tweet events:

- OntoDSumm [23]: Uses ontology-based techniques to improve summary quality by incorporating domain-specific knowledge.
- OntoRealSumm [51]: Integrates real-world ontologies for more contextually relevant summaries.
- SOM+GSOM [21]: A clustering-based approach that combines self-organizing maps (SOMs) with Generalized SOM (GSOM) to create clusters of text, facilitating the extraction of meaningful content for summaries.
- SOM+SOM [21]: Another variant of SOM. This method uses multiple layers of SOM to better model complex relationships between text segments.
- MOOTweetSumm [22]: A multi-objective extractive summarization approach for tackling the microblog summarization challenge.
- Stream-3GANSumm [26]: Uses multi-view data representations and a triple-generative adversarial network (3GAN) variant for real-time summarization.
- Rel-Cov-Red [52]: A multi-objective pruning approach extracts key concepts from opinionated content, using manifold learning and clustering to optimize relevancy, redundancy, and coverage.
- Rel-Red-Cov [52]: An extension of Rel-Cov-Red, it refines the coverage reduction process by balancing relevance and redundancy.
- ATSumm [29]: It is an abstractive disaster tweet summarization approach that enhances summary quality by integrating auxiliary information.
- IKDSumm [33]: It is an extractive summarization approach for disaster-related tweets, enhancing summary quality by leveraging key phrases and BERT.

These methods provide a basis for evaluating the performance of our model across different summarization strategies.

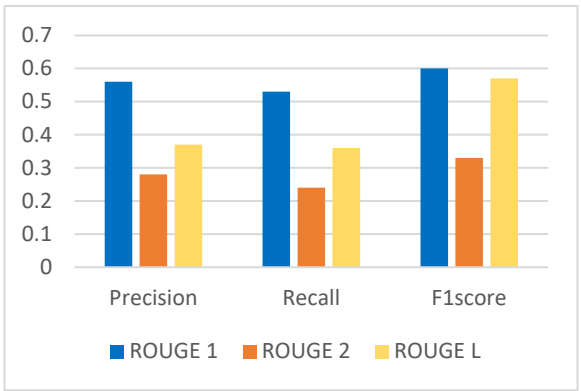
### *4.2. Evaluation of Proposed Framework Performance*

Table 3 provides a detailed comparative analysis of performance metrics, including recall, precision, and F1 score, for the proposed hybrid BERT-LSTM framework across three datasets: Hagupit, SHShoot, and Hyderabad Blast. The evaluation is based on ROUGE metrics [53], specifically ROUGE-1, ROUGE-2, and ROUGE-L, to measure the effectiveness of the proposed framework. The hybrid BERT-LSTM demonstrates superior performance, achieving higher recall, precision, and F1 scores across most of the datasets compared to conventional extractive summarization methods.

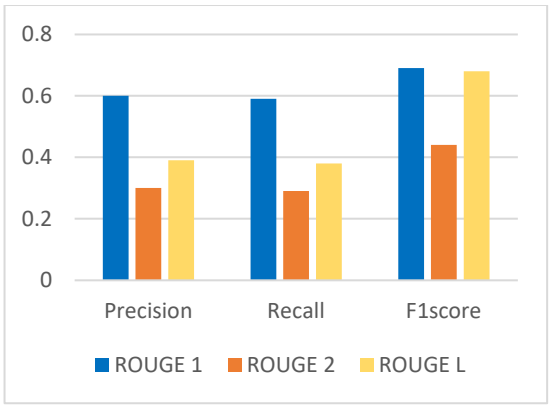
**Table 3.** Performance comparison of the proposed framework.

| Dataset | *ROUGE* | Precision | Recall | F1score |
|---|---|---|---|---|
| **Hagupit** | *ROUGE* 1 | 0.56 | 0.53 | 0.60 |
| | *ROUGE* 2 | 0.28 | 0.24 | 0.33 |
| | *ROUGE L* | 0.37 | 0.36 | 0.57 |
| **SHShoot** | *ROUGE* 1 | 0.60 | 0.59 | 0.69 |
| | *ROUGE* 2 | 0.30 | 0.29 | 0.44 |
| | *ROUGE L* | 0.39 | 0.38 | 0.68 |
| **Hyderabad Blast** | *ROUGE* 1 | 0.55 | 0.53 | 0.70 |
| | *ROUGE* 2 | 0.29 | 0.26 | 0.43 |
| | *ROUGE L* | 0.38 | 0.36 | 0.66 |

Figure 4 reflects the hybrid BERT-LSTM framework's ability to generate high-quality summaries across diverse datasets. Specifically, the framework excels in capturing essential information, as reflected by the high ROUGE-1 scores across all datasets (0.60 for Hagupit, 0.69 for SHShoot, and 0.70 for Hyderabad Blast). Its ability to preserve structural coherence is evidenced by competitive ROUGE-L scores, particularly in datasets like SHShoot and Hyderabad Blast, with F1 scores of 0.68 and 0.66, respectively. Meanwhile, the ROUGE-2 scores (0.33 for Hagupit, 0.44 for SHShoot, and 0.43 for Hyderabad Blast) demonstrate moderate retention of contextual relationships. The framework also shows a favorable balance between precision and recall, with recall generally higher than precision, signifying a focus on capturing comprehensive content. This is particularly evident in the Hyderabad Blast dataset, which achieved the highest recall score (0.70 for ROUGE-1). Overall, the hybrid BERT-LSTM framework proves its efficacy in generating high-quality summaries across diverse datasets, showcasing its adaptability and superiority.

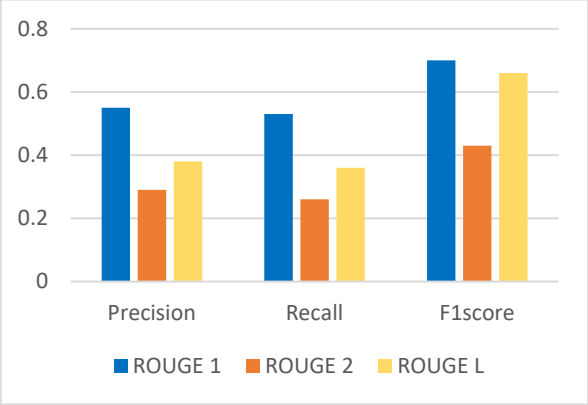



(**a**) Hagupit




(**b**) SHShoot

**Figure 4.** *Cont.*




(**c**) Hyderabad Blast

**Figure 4.** Performance comparison of the proposed framework.

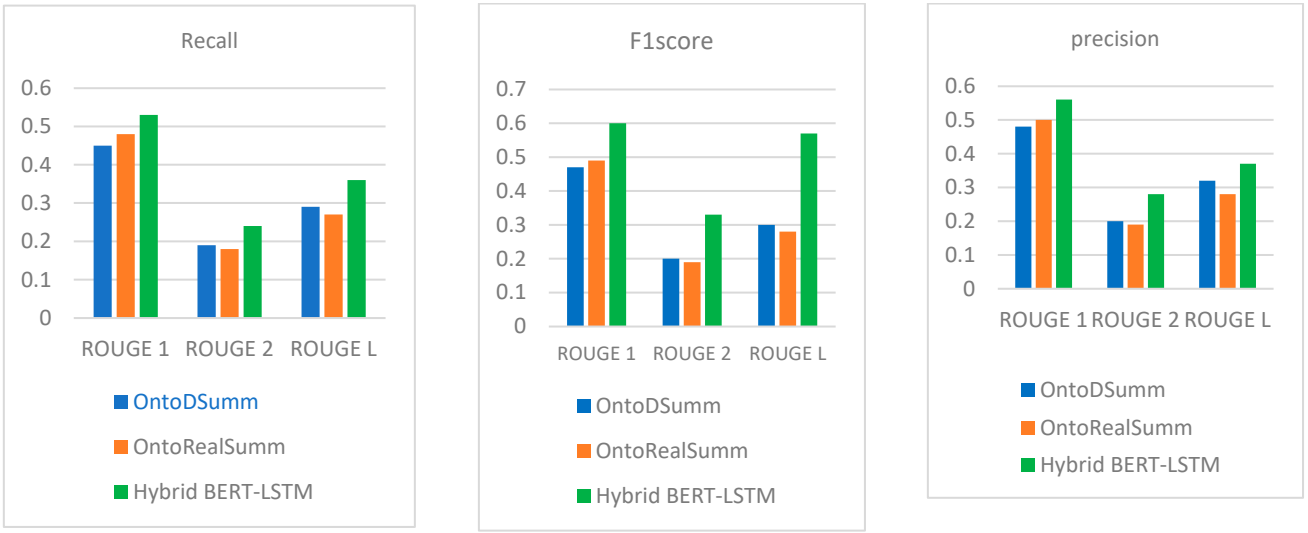
### *4.3. Discussions*

Table 4 and Figure 5 present a comparative analysis of the proposed BERT-LSTM framework performance against two established summarization techniques, namely OntoDSumm and OntoRealSumm, across three distinct disaster datasets: Hagupit, SHShoot, and Hyderabad Blast. Each model's performance was assessed using several metrics: F1 score, recall, and precision, evaluated through ROUGE 1, ROUGE 2, and ROUGE L measures to reveal significant insights into the effectiveness of the proposed framework. The hybrid BERT-LSTM framework demonstrates superior performance, achieving higher recall, precision, and F1 scores across all datasets compared to conventional extractive summarization methods. These scores underscore BERT-LSTM's ability to capture contextual dependencies effectively, enhancing both precision and recall compared to other summarization techniques.

The results for the Hagupit dataset reveal significant performance of the proposed framework differences among the summarization methods in all metrics, particularly a ROUGE-1 score of 0.60, a ROUGE-2 score of 0.33, and a ROUGE-L score of 0.57, indicating better overall summary quality. In contrast, OntoDSumm produced a ROUGE-1 score of 0.47, a ROUGE-2 score of 0.20, and a ROUGE-L score of 0.30, while OntoRealSumm recorded slightly lower scores, with ROUGE-1 at 0.49, ROUGE-2 at 0.19, and ROUGE-L at 0.28. The BERT-LSTM consistently outperforms both OntoDSumm and OntoRealSumm across all datasets and metrics. Again, performance was similarly high on the SHShoot dataset, where the BERT-LSTM framework recorded a ROUGE-1 score of 0.69, a ROUGE-2 score of 0.44, and a ROUGE-L score of 0.68. The OntoDSumm method yielded ROUGE-1, ROUGE-2, and ROUGE-L scores of 0.54, 0.23, and 0.29, respectively, while OntoRealSumm produced scores of 0.51 (ROUGE-1), 0.20 (ROUGE-2), and 0.29 (ROUGE-L). The performance metrics for the Hyderabad Blast dataset support the findings from the previous datasets. The BERT-LSTM framework achieved a ROUGE-1 score of 0.70, a ROUGE-2 score of 0.43, and a ROUGE-L score of 0.66. On the other hand, OntoDSumm and OntoRealSumm exhibited ROUGE-1 scores of 0.47 and 0.47, ROUGE-2 scores of 0.19 and 0.18, and ROUGE-L scores of 0.28 for both methods, respectively. Furthermore, the proposed framework achieved perfect recall and precision across all ROUGE metrics, indicating its strong ability to capture all relevant content from the reference summaries. The consistency of the BERT-LSTM's performance across all datasets further validates its effectiveness. This indicates that integrating BERT with LSTM not only enhances the quality of summaries but also improves their coherence and relevance.

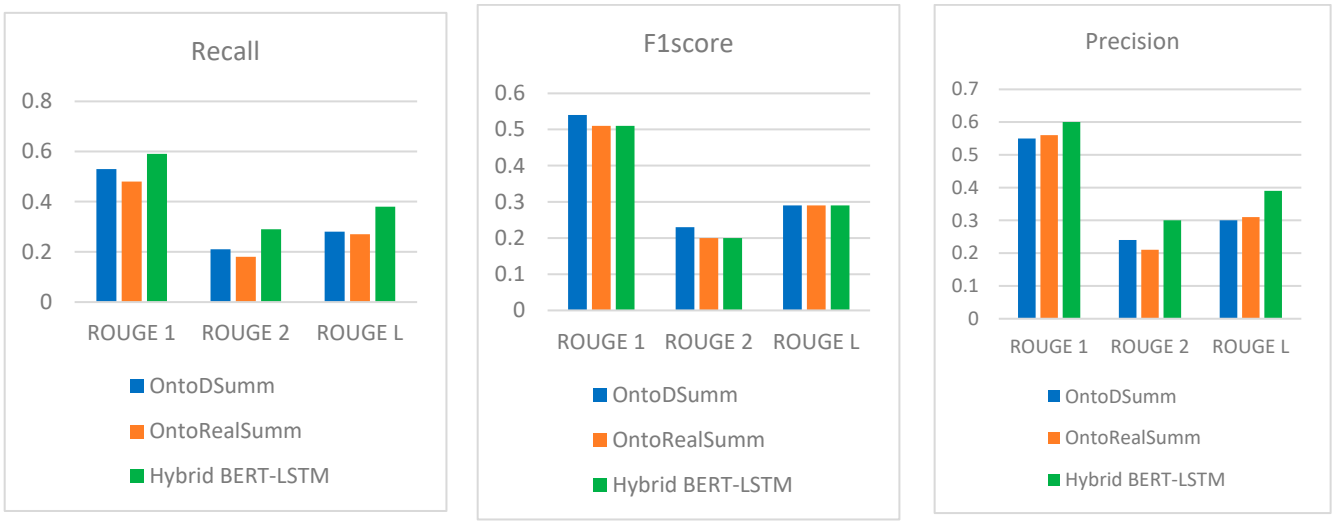
**Table 4.** Performance comparison of the proposed framework and existing methods.

| Dataset | Technique | Measures | ROUGE 1 | ROUGE 2 | ROUGE L |
|---|---|---|---|---|---|
| **Hagupit** | OntoDSumm [23] | F1 score | 0.47 | 0.20 | 0.30 |
| | | Recall | 0.45 | 0.19 | 0.29 |
| | | Precision | 0.48 | 0.20 | 0.32 |
| | OntoRealSumm | F1 score | 0.49 | 0.19 | 0.28 |
| | | Recall | 0.48 | 0.18 | 0.27 |
| | | Precision | 0.50 | 0.19 | 0.28 |
| | Hybrid BERT-LSTM | F1 score | 0.60 | 0.33 | 0.57 |
| | | Recall | 0.53 | 0.24 | 0.36 |
| | | Precision | 0.56 | 0.28 | 0.37 |
| **SHShoot** | OntoDSumm | F1 score | 0.54 | 0.23 | 0.29 |
| | | Recall | 0.53 | 0.21 | 0.28 |
| | | Precision | 0.55 | 0.24 | 0.30 |
| | OntoRealSumm | F1 score | 0.51 | 0.20 | 0.29 |
| | | Recall | 0.48 | 0.18 | 0.27 |
| | | Precision | 0.56 | 0.21 | 0.31 |
| | Hybrid BERT-LSTM | F1 score | 0.69 | 0.44 | 0.68 |
| | | Recall | 0.59 | 0.29 | 0.38 |
| | | Precision | 0.60 | 0.30 | 0.39 |
| **Hyderabad Blast** | OntoDSumm | F1 score | 0.47 | 0.19 | 0.28 |
| | | Recall | 0.48 | 0.19 | 0.28 |
| | | Precision | 0.48 | 0.20 | 0.29 |
| | OntoRealSumm | F1 score | 0.47 | 0.18 | 0.28 |
| | | Recall | 0.44 | 0.17 | 0.27 |
| | | Precision | 0.49 | 0.19 | 0.29 |
| | Hybrid BERT-LSTM | F1 score | 0.70 | 0.43 | 0.66 |
| | | Recall | 0.53 | 0.26 | 0.36 |
| | | Precision | 0.55 | 0.29 | 0.38 |

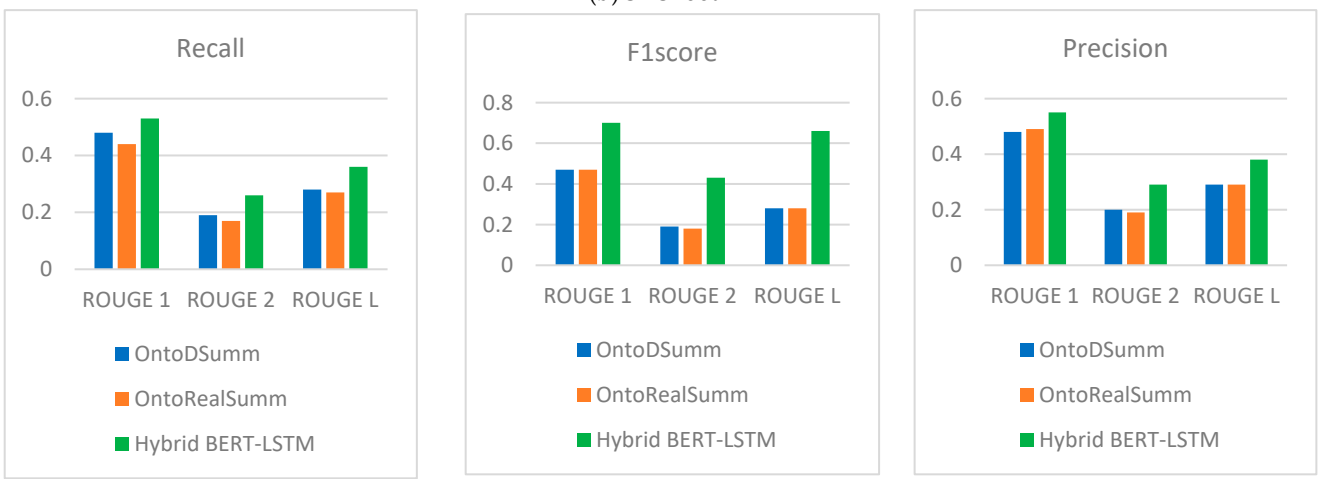



(**a**) Hagupit

**Figure 5.** *Cont.*




(**b**) SHShoot




(**c**) Hyderabad Blast

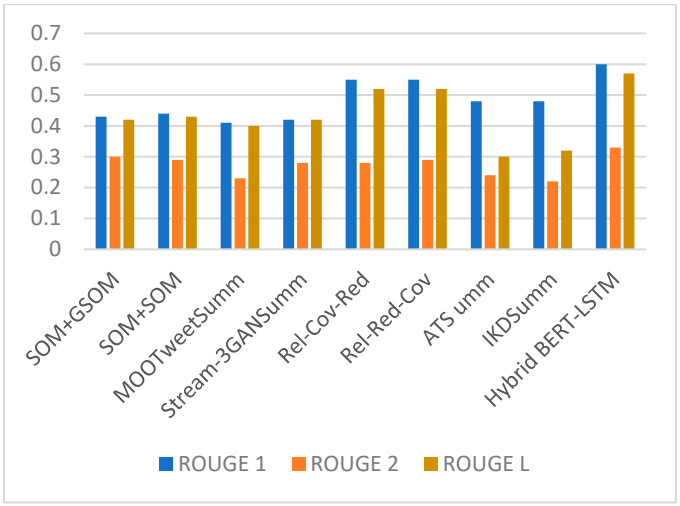
**Figure 5.** Performance comparison of the proposed and existing framework.

In Table 5 and Figure 6, which indicate the performance of F1 scores compared to the existing methods, the results show ROUGE − 1, ROUGE − 2, and ROUGE − L. These metrics suggest that hybrid BERT-LSTM outperforms the existing method in terms of extractive summarization quality, particularly in capturing the essence and key elements of the original text across various evaluation dimensions like n-gram overlap and longest common subsequence. The proposed hybrid BERT-LSTM framework demonstrates superior performance compared to the existing approaches. Our proposed method achieves higher F1 scores across most of the datasets. In particular, the hybrid BERT-LSTM framework consistently outperformed its counterparts, achieving the highest F1 scores, with values of 0.60 and 0.70 for ROUGE-1 across the Hagupit and Hyderabad Blast datasets, respectively. This model's performance highlights its capability to effectively capture both contextual relevance and linguistic coherence, crucial for summarizing complex narratives often found in disaster-related content. Hybrid BERT-LSTM achieves F1 scores of 0.70 (ROUGE-1) and 0.66 (ROUGE-L) on the Hyderabad Blast dataset, indicating its robustness in capturing keywords and maintaining sequence coherence. Hybrid BERT-LSTM outperforms other models in terms of the F1 score across the most datasets, demonstrating high scores, particularly on ROUGE-1 and ROUGE-L metrics.

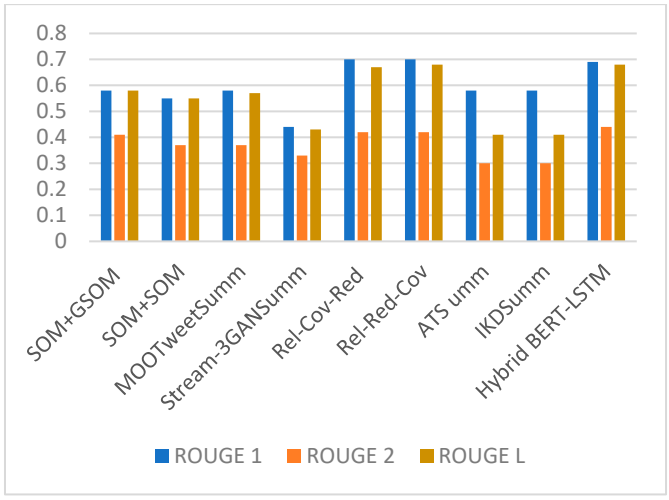
**Table 5.** Performance comparison of the existing framework.

| Dataset | Technique | ROUGE 1 F1 Score | ROUGE 2 F1 Score | ROUGE L F1 Score |
|---|---|---|---|---|
| **Hagupit** | SOM+GSOM | 0.43 | 0.30 | 0.42 |
| | SOM+SOM | 0.44 | 0.29 | 0.43 |
| | MOOTweetSumm | 0.41 | 0.23 | 0.40 |
| | Stream-3GANSumm | 0.42 | 0.28 | 0.42 |
| | Rel-Cov-Red | 0.55 | 0.28 | 0.52 |
| | Rel-Red-Cov | 0.55 | 0.29 | 0.52 |
| | ATSumm | 0.48 | 0.24 | 0.30 |
| | IKDSumm | 0.48 | 0.22 | 0.32 |
| | Hybrid BERT-LSTM | 0.60 | 0.33 | 0.57 |
| **SHShoot** | SOM+GSOM | 0.58 | 0.41 | 0.58 |
| | SOM+SOM | 0.55 | 0.37 | 0.55 |
| | MOOTweetSumm | 0.58 | 037 | 0.57 |
| | Stream-3GANSumm | 0.44 | 0.33 | 0.43 |
| | Rel-Cov-Red | 0.70 | 0.42 | 0.67 |
| | Rel-Red-Cov | 0.70 | 0.42 | 0.68 |
| | ATSumm | 0.58 | 0.30 | 0.41 |
| | IKDSumm | 0.58 | 0.30 | 0.41 |
| | Hybrid BERT-LSTM | 0.69 | 0.44 | 0.68 |
| **Hyderabad Blast** | SOM+GSOM | 0.53 | 0.41 | 0.53 |
| | SOM+SOM | 0.52 | 0.39 | 0.51 |
| | MOOTweetSumm | 0.52 | 0.35 | 0.51 |
| | Stream-3GANSumm | 0.44 | 0.31 | 0.43 |
| | Rel-Cov-Red | 0.54 | 0.28 | 0.52 |
| | Rel-Red-Cov | 0.55 | 0.29 | 0.52 |
| | ATS umm | 0.52 | 0.28 | 0.34 |
| | IKDSumm | 0.52 | 0.28 | 0.34 |
| | Hybrid BERT-LSTM | 0.70 | 0.43 | 0.66 |

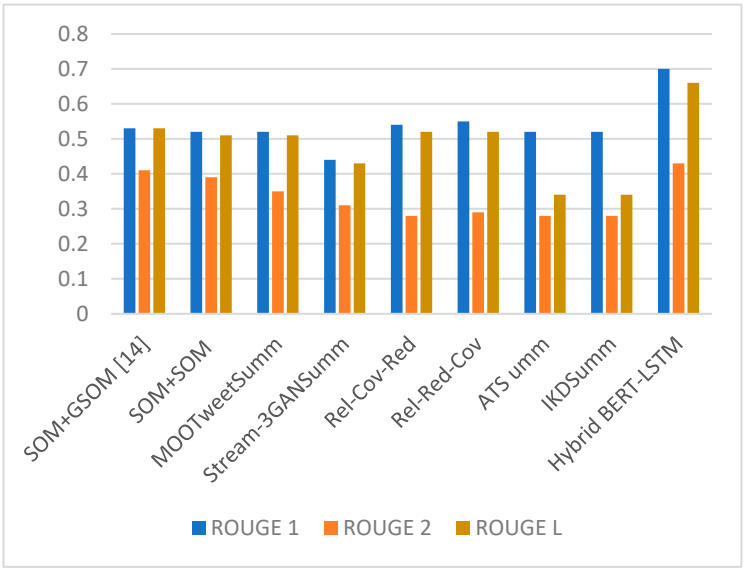
Models such as Rel-Cov-Red and Rel-Red-Cov performed well, especially in the SHShoot dataset, achieving ROUGE-1 scores of up to 0.70. In contrast, SOM-based models, including SOM+GSOM and SOM+SOM, showcased moderate performance, with F1 scores around 0.43 to 0.58 for ROUGE-1. However, these models fell short in generating fluent summaries, as evidenced by their lower ROUGE-2 scores, suggesting difficulties in preserving the linguistic structures necessary for coherent summaries. Furthermore, their performance in ROUGE-2 scores indicates potential challenges in maintaining the complexity and depth of the original narratives. This performance gap highlights the importance of not only identifying relevant content but also ensuring the grammatical and contextual integrity of the summaries, which is particularly challenging in complex datasets. Overall, the results underscore the efficiency of the proposed hybrid BERT-LSTM model, indicating a clear trend toward deep learning methodologies that leverage advanced contextual representations to enhance summary quality.




(**a**) Hagupit




(**b**) SHShoot




(**c**) Hyderabad Blast

**Figure 6.** Performance comparison between the proposed Hybrid BERT-LSTM framework and existing methods.

### *4.4. Overall Observations*

The overall observation from the discussions above reveals that while several summarization methods offer valuable insights for processing disaster-related tweet data, they each have unique strengths and limitations that impact their effectiveness in different contexts. Techniques like Rel-Cov-Red and Rel-Red-Cov methods also perform relatively well, particularly in the SHShoot dataset, with high ROUGE-1 F1 scores of 0.70, underscoring their capability for thorough information coverage. However, these models sometimes compromise on fluency, as suggested by their slightly lower ROUGE-L scores in other datasets, which impacts the overall readability of the summaries. The OntoDSumm, OntoRealSumm, ATS umm, and IKDSumm models generally yield moderate scores across datasets, with F1 scores ranging between 0.47 and 0.58, indicating that while they identify relevant content, they often lack the precision or cohesion observed in other models. Cluster-based models like SOM+GSOM and SOM+SOM are slightly more effective in identifying thematic information, especially notable in the SHShoot dataset with ROUGE-2 and ROUGE-L F1 scores of 0.41 and 0.58, respectively. However, these models struggle with maintaining coherent sentence structures, particularly in more complex datasets. Stream-3GANSumm,

optimized for real-time processing, delivers moderate performance, with F1 scores hovering around 0.42–0.44 across the datasets, indicating that while it is fast, it may miss nuanced information, especially when inputs are dense or complex.

To address these issues, the hybrid BERT-LSTM was proposed. This model captures the strength both of BERT and LSTM toward well-inherent relationships between text segments through deep contextual understanding and fluency in summarization by the structure of LSTM, which is effective in sequential information processing. This combination will guarantee summaries are not only contextually rich but coherent and concise. This model offers a strong summarization solution for the very heterogeneous and complex input text, eliminating redundancy and furthering coherence in output. The hybrid BERT-LSTM framework consistently demonstrates the highest effectiveness, achieving the highest ROUGE-1 F1 scores across all datasets (0.60 for Hagupit, 0.69 for SHShoot, and 0.70 for Hyderabad Blast). This superior performance reflects its strong ability to capture relevant content and maintain coherence, especially when compared to other models.

## 5. Conclusions

This study introduced an innovative framework for automated short text summarization on social media, integrating a novel BERT with a transformer-based encoder–decoder and an attention mechanism. The approach was specifically designed to handle the challenges of summarizing brief, informal, and noisy text. To evaluate its effectiveness, we tested the framework on three datasets (Hagupit, SHShoot, and Hyderabad Blast) using ROUGE scores as key performance metrics. The results demonstrated that our method outperforms existing techniques in generating high-quality summaries. Our findings highlight the framework's ability to produce concise, coherent summaries while preserving essential information. The combination of BERT's advanced language comprehension and LSTM's sequential processing contributed to generating more accurate and cohesive summaries, reinforcing its effectiveness in social media text summarization. Despite these promising results, further refinements are necessary, particularly in optimizing the balance between recall and precision. Future research should aim to refine precision while maintaining recall, enhance model adaptability, and incorporate additional evaluation metrics such as CIDEr and SPICE. Features like real-time summarization, personalized summaries, and domain-specific adaptations could further improve the framework's practical impact. To achieve a better balance between precision and recall, multi-objective optimization techniques, such as Pareto optimization and reinforcement learning, could be utilized to adjust model parameters in real time. Adaptive loss functions could also be implemented to optimize performance based on the specific characteristics of the input data. For time-sensitive applications, implementing streaming data processing and token-level parallelism could enable faster summary generation. User profiling could personalize summaries, enhancing relevance, while cross-lingual transfer learning and training on diverse datasets would increase the model's versatility. Furthermore, using evaluation metrics like BLEU, TER, or METEOR would offer a more comprehensive assessment of the summaries' context, fluency, and diversity.

**Author Contributions:** Conceptualization, F.A.G., M.C.P. and R.A.; methodology, F.A.G., H.M.A., M.C.P. and R.A.; software, F.A.G. and H.M.A.; formal analysis, F.A.G., M.C.P. and R.A.; investigation, R.A. and M.C.P.; resources, F.A.G. and R.A.; data curation, F.A.G., H.M.A. and R.A.; writing—original draft preparation, F.A.G., R.A. and M.C.P.; writing—review and editing, F.A.G., R.A. and M.C.P.; visualization, F.A.G. and H.M.A.; supervision, M.C.P. and R.A.; project administration, M.C.P.; funding acquisition, F.A.G. and R.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This paper received no external funding.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are contained within the article.

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

## References

1. Ghanem, F.A.; Padma, M.C.; Alkhatib, R. Automatic Short Text Summarization Techniques in Social Media Platforms. *Future Internet* **2023**, *15*, 311. [CrossRef]
2. Supriyono; Wibawa, A.P.; Suyono; Kurniawan, F. A survey of text summarization: Techniques, evaluation and challenges. *Nat. Lang. Process. J.* **2024**, *7*, 100070. [CrossRef]
3. Widyassari, A.P.; Rustad, S.; Shidik, G.F.; Noersasongko, E.; Syukur, A.; Affandy, A.; Setiadi, D.R.I.M. Review of automatic text summarization techniques & methods. *J. King Saud Univ.—Comput. Inf. Sci.* **2022**, *34*, 1029–1046. [CrossRef]
4. Shakil, H.; Farooq, A.; Kalita, J. Abstractive text summarization: State of the art, challenges, and improvements. *Neurocomputing* **2024**, *603*, 128255. [CrossRef]
5. Yadav, A.K.; Ranvijay; Yadav, R.S.; Maurya, A.K. State-of-the-art approach to extractive text summarization: A comprehensive review. *Multimed. Tools Appl.* **2023**, *82*, 29135–29197. [CrossRef]
6. Baykara, B.; Güngör, T. Abstractive text summarization and new large-scale datasets for agglutinative languages Turkish and Hungarian. *Lang. Resour. Eval.* **2022**, *56*, 973–1007. [CrossRef]
7. Mohiuddin, K.; Alam, M.A.; Alam, M.M.; Welke, P.; Martin, M.; Lehmann, J.; Vahdati, S. Retention is All You Need. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management, Birmingham, UK, 21–25 October 2023*; ACM: New York, NY, USA, 2023; pp. 4752–4758.
8. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the NAACL HLT 2019–2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies—Proceedings of the Conference, Mineapolis, MN, USA, 2–7 June 2019; Volume 1, pp. 4171–4186. Available online: http://arxiv.org/abs/1810.04805 (accessed on 8 November 2024).
9. Saleh, M.E.; Wazery, Y.M.; Ali, A.A. A systematic literature review of deep learning-based text summarization: Techniques, input representation, training strategies, mechanisms, datasets, evaluation, and challenges. *Expert Syst. Appl.* **2024**, *252*, 124153. [CrossRef]
10. Liu, Y.; Lapata, M. Text summarization with pretrained encoders. In *Proceedings of the EMNLP-IJCNLP 2019–2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference, Hong Kong, China, 10–13 November 2019*; Association for Computational Linguistics: Stroudsburg, PA, USA, 2019; pp. 3730–3740.
11. See, A.; Liu, P.J.; Manning, C.D. Get to the point: Summarization with pointer-generator networks. In *Proceedings of the ACL 2017 55th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers), Vancouver, BC, Canada, 30 July–4 August 2017*; Association for Computational Linguistics: Stroudsburg, PA, USA, 2017; Volume 1, pp. 1073–1083. [CrossRef]
12. Zhang, H.; Cai, J.; Xu, J.; Wang, J. Pretraining-Based Natural Language Generation for Text Summarization. In *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL), Hong Kong, China, 3–4 November 2019*; Association for Computational Linguistics: Stroudsburg, PA, USA, 2019; pp. 789–797.
13. El-Kassas, W.S.; Salama, C.R.; Rafea, A.A.; Mohamed, H.K. Automatic text summarization: A comprehensive survey. *Expert Syst. Appl.* **2021**, *165*, 113679. [CrossRef]
14. Sharma, G.; Sharma, D. Automatic Text Summarization Methods: A Comprehensive Review. *SN Comput. Sci.* **2022**, *4*, 33. [CrossRef]
15. Balaji, N.; Deepa, K.; Bhavatarini, N.; Megha, N.; Sunil, K.P.; Shikah, R.A. Text Summarization using NLP Technique. In Proceedings of the 2022 International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), Shivamogga, Karnataka, India, 14–15 October 2022; pp. 30–35.
16. Anand, D.; Wagh, R. Effective deep learning approaches for summarization of legal texts. *J. King Saud Univ.—Comput. Inf. Sci.* **2022**, *34*, 2141–2150. [CrossRef]
17. Jain, A.; Arora, A.; Morato, J.; Yadav, D.; Kumar, K.V. Automatic Text Summarization for Hindi Using Real Coded Genetic Algorithm. *Appl. Sci.* **2022**, *12*, 6584. [CrossRef]
18. Ghanem, F.A.; Padma, M.C.; Abdulwahab, H.M.; Alkhatib, R. Novel Genetic Optimization Techniques for Accurate Social Media Data Summarization and Classification Using Deep Learning Models. *Technologies* **2024**, *12*, 199. [CrossRef]

19. Ramachandran, D.; Ramasubramanian, P. Event detection from Twitter—a survey. *Int. J. Web Inf. Syst.* **2018**, *14*, 262–280. [CrossRef]
20. Panchendrarajan, R.; Hsu, W.; Li Lee, M. Emotion-Aware Event Summarization in Microblogs. In *Companion Proceedings of the Web Conference, Ljubljana, Slovenia, 19–23 April 2021*; ACM: New York, NY, USA, 2021; pp. 486–494.
21. Saini, N.; Saha, S.; Mansoori, S.; Bhattacharyya, P. Fusion of self-organizing map and granular self-organizing map for microblog summarization. *Soft Comput.* **2020**, *24*, 18699–18711. [CrossRef]
22. Saini, N.; Saha, S.; Bhattacharyya, P. Multiobjective-Based Approach for Microblog Summarization. *IEEE Trans. Comput. Soc. Syst.* **2019**, *6*, 1219–1231. [CrossRef]
23. Garg, P.K.; Chakraborty, R.; Dandapat, S.K. OntoDSumm: Ontology-Based Tweet Summarization for Disaster Events. *IEEE Trans. Comput. Soc. Syst.* **2024**, *11*, 2724–2739. [CrossRef]
24. Huang, Y.; Shen, C.; Li, T. Event summarization for sports games using twitter streams. *World Wide Web* **2018**, *21*, 609–627. [CrossRef]
25. Goyal, P.; Kaushik, P.; Gupta, P.; Vashisth, D.; Agarwal, S.; Goyal, N. Multilevel Event Detection, Storyline Generation, and Summarization for Tweet Streams. *IEEE Trans. Comput. Soc. Syst.* **2020**, *7*, 8–23. [CrossRef]
26. Paul, D.; Rana, S.; Saha, S.; Mathew, J. Online Summarization of Microblog Data: An Aid in Handling Disaster Situations. *IEEE Trans. Comput. Soc. Syst.* **2024**, *11*, 4029–4039. [CrossRef]
27. Garg, P.K.; Chakraborty, R.; Dandapat, S.K. PORTRAIT: A hybrid aPproach tO cReate extractive ground-TRuth summAry for dIsaster evenT. *ACM Trans. Web* **2025**, *19*, 1–36. Available online: http://arxiv.org/abs/2305.11536 (accessed on 4 April 2024). [CrossRef]
28. Madichetty, S.; Muthukumarasamy, S. Detection of situational information from Twitter during disaster using deep learning models. *Sādhanā* **2020**, *45*, 270. [CrossRef]
29. Garg, P.K.; Chakraborty, R.; Dandapat, S.K. ATSumm: Auxiliary information enhanced approach for abstractive disaster Tweet Summarization with sparse training data. *arXiv* **2024**, arXiv:2405.06541. [CrossRef]
30. Lin, C.; Ouyang, Z.; Wang, X.; Li, H.; Huang, Z. Preserve Integrity in Realtime Event Summarization. *ACM Trans. Knowl. Discov. Data* **2021**, *15*, 1–29. [CrossRef]
31. Nguyen, T.H.; Rudra, K. Towards an Interpretable Approach to Classify and Summarize Crisis Events from Microblogs. In *Proceedings of the ACM Web Conference, Lyon, France, 25–29 April 2022*; ACM: New York, NY, USA, 2022; pp. 3641–3650.
32. Bansal, D.; Saini, N.; Saha, S. DCBRTS: A Classification-Summarization Approach for Evolving Tweet Streams in Multiobjective Optimization Framework. *IEEE Access* **2021**, *9*, 148325–148338. [CrossRef]
33. Garg, P.K.; Chakraborty, R.; Gupta, S.; Dandapat, S.K. IKDSumm: Incorporating key-phrases into BERT for extractive disaster tweet summarization. *Comput. Speech Lang.* **2024**, *87*, 101649. [CrossRef]
34. Rudra, K.; Goyal, P.; Ganguly, N.; Imran, M.; Mitra, P. Summarizing Situational Tweets in Crisis Scenarios: An Extractive-Abstractive Approach. *IEEE Trans. Comput. Soc. Syst.* **2019**, *6*, 981–993. [CrossRef]
35. Sen, A.; Rudra, K.; Ghosh, S. Extracting situational information from microblogs during disaster events: A classification-summarization approach. In Proceedings of the 2015 7th International Conference on Communication Systems and Networks (COMSNETS), Bangalore, India, 6–10 January 2015; pp. 1–6.
36. Abdel-Salam, S.; Rafea, A. Performance Study on Extractive Text Summarization Using BERT Models. *Information* **2022**, *13*, 67. [CrossRef]
37. La Quatra, M.; Cagliero, L. BART-IT: An Efficient Sequence-to-Sequence Model for Italian Text Summarization. *Future Internet* **2022**, *15*, 15. [CrossRef]
38. Zhao, Y.; Huang, S.; Zhou, D.; Ding, Z.; Wang, F.; Nian, A. CNsum:Automatic Summarization for Chinese News Text. *arXiv* **2025**, arXiv:2502.19723. [CrossRef]
39. Divya, S.; Sripriya, N.; Andrew, J.; Mazzara, M. Unified extractive-abstractive summarization: A hybrid approach utilizing BERT and transformer models for enhanced document summarization. *PeerJ Comput. Sci.* **2024**, *10*, e2424.
40. Abadi, V.N.M.; Ghasemian, F. Enhancing Persian text summarization through a three-phase fine-tuning and reinforcement learning approach with the mT5 transformer model. *Sci. Rep.* **2025**, *15*, 80. [CrossRef] [PubMed]
41. Papagiannopoulou, A.; Angeli, C. Encoder-Decoder Transformers for Textual Summaries on Social Media Content. *Autom. Control Intell. Syst.* **2024**, *12*, 48–59. [CrossRef]
42. Kherwa, P.; Arora, J.; Sharma, T.; Gupta, D.; Juneja, S.; Muhammad, G.; Nauman, A. Contextual embedded text summarizer system: A hybrid approach. *Expert Syst.* **2025**, *42*, e13733. [CrossRef]
43. Toprak, A.; Turan, M. Enhanced automatic abstractive document summarization using transformers and sentence grouping. *J. Supercomput.* **2025**, *81*, 1–30. [CrossRef]
44. Murugaraj, K.; Lamsiyah, S.; Schommer, C. Abstractive Summarization of Historical Documents: A New Dataset and Novel Method using a Domain-Specific Pretrained Model. *IEEE Access* **2025**, *13*, 10918–10932. [CrossRef]

45. Zhang, H.; Liu, X.; Zhang, J. SummIt: Iterative Text Summarization via ChatGPT. In *Proceedings of the Findings of the Association for Computational Linguistics: EMNLP 2023, Singapore, 6–10 December 2023*; Association for Computational Linguistics: Stroudsburg, PA, USA, 2023; pp. 10644–10657.
46. Ghanem, F.A.; Padma, M.C.; Alkhatib, R. Elevating the Precision of Summarization for Short Text in Social Media using Preprocessing Techniques. In Proceedings of the 2023 IEEE International Conference on High Performance Computing & Communications, Data Science Systems, Smart City & Dependability in Sensor, Cloud & Big Data Systems Application (HPCC/DSS/SmartCity/DependSys), Melbourne, Australia, 17–21 December 2023; pp. 408–416.
47. Bano, S.; Khalid, S.; Tairan, N.M.; Shah, H.; Khattak, H.A. Summarization of scholarly articles using BERT and BiGRU: Deep learning-based extractive approach. *J. King Saud Univ.—Comput. Inf. Sci.* **2023**, *35*, 101739. [CrossRef]
48. Bano, S.; Khalid, S. BERT-based Extractive Text Summarization of Scholarly Articles: A Novel Architecture. In Proceedings of the 2022 International Conference on Artificial Intelligence of Things (ICAIoT), Istanbul, Turkey, 29–30 December 2022; pp. 1–5.
49. Bani-Almarjeh, M.; Kurdy, M.-B. Arabic abstractive text summarization using RNN-based and transformer-based architectures. *Inf. Process. Manag.* **2023**, *60*, 103227. [CrossRef]
50. Gray, D.; Bowes, D.; Davey, N.; Sun, Y.; Christianson, B. Further thoughts on precision. *IET Semin. Dig.* **2011**, *2011*, 129–133. [CrossRef]
51. Garg, P.K.; Chakraborty, R.; Dandapat, S.K. OntoRealSumm: Ontology based Real-Time Tweet Summarization. *arXiv* **2022**, arXiv:2201.06545. [CrossRef]
52. Gudakahriz, S.J.; Moghadam, A.M.E.; Mahmoudi, F. Opinion texts summarization based on texts concepts with multi-objective pruning approach. *J. Supercomput.* **2023**, *79*, 5013–5036. [CrossRef]
53. Lin, C.-Y. ROUGE: A Package for Automatic Evaluation of Summaries Chin-Yew. *Text Summ. Branches Out* **2004**, 74–81. Available online: https://aclanthology.org/W04-1013/ (accessed on 4 April 2024).


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