AVE Trends in Intelligent Computing Systems

Enhancing Abstractive Text Summarization with Neural Network-Based Techniques

S. Hemadharsana[1,*], N. Madhumitha[2], R. Vinoth[3], R. Vanitha[4], G. Brintha[5], Adithi Venkatakrishnan[6]

[1]Department of Computer Science and Engineering (Cyber Security), Saveetha Engineering College, Chennai, Tamil Nadu, India.
[2]Department of Computer Science and Engineering, Saveetha Engineering College, Chennai, Tamil Nadu, India.
[3]Department of Artificial Intelligence and Machine Learning, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
[4]Department of Computer Science and Engineering, KCG College of Technology, Chennai, Tamil Nadu, India.
[5]Department of Artificial Intelligence and Data Science, Francis Xavier Engineering College, Tirunelveli, Tamil Nadu, India.
[6]Department of Information Technology and Management, University of Texas at Dallas, Richardson, Texas, United States of America.
hemadharsana@gmail.com[1], madhumitha248@gmail.com[2], vinothr2@srmist.edu.in[3], vanitha.cse@kcgcollege.com[4], gbrinthanagarajan@gmail.com[5], dal076222@utdallas.edu[6]

*Corresponding author

Abstract: Long Short-Term Memory (LSTM) networks in an encoder–decoder architecture are used to test abstractive text summarisation for news articles. The InShorts News Article Dataset is used to train and evaluate several deep learning configurations, including regular LSTM and BiGRU models, LSTMs with pre-trained word embeddings such as GloVe and Word2Vec, and attention-integrated architectures. While researchers examined various performance measures, BLEU scores were the most essential for summarising quality. The LSTM encoder-decoder with pre-trained Word2Vec embeddings and an attention layer achieved the highest BLEU score of 0.7481. This illustrates that attention approaches significantly affect how well the model catches contextual dependencies and prioritises key input content. Pre-trained embeddings improve semantic understanding and summarisation. The findings indicate that encoder-decoder designs and attention-based improvements are crucial for abstractive summarisation. With copious data, efficient text summarisation is crucial for better information retrieval, faster content consumption, and user understanding in the digital age. Neural architecture comparison, hyperparameter optimization, transformer-based models, and new embedding strategies to improve summarisation accuracy and resilience are future research topics.

Keywords: Hyperparameter Optimisation; Abundant Data; Abstractive Summarisation; Attention Mechanisms; Neural Architecture; Encoder–Decoder Architecture; Transformer-Based Models.

Cite as: S. Hemadharsana, N. Madhumitha, R. Vinoth, R. Vanitha, G. Brintha, and A. Venkatakrishnan, "Enhancing Abstractive Text Summarization with Neural Network-Based Techniques," *AVE Trends in Intelligent Computing Systems*, vol. 2, no. 4, pp. 184–195, 2025.

Journal Homepage: https://www.avepubs.com/user/journals/details/ATICS

Received on: 13/12/2024, **Revised on:** 03/04/2025, **Accepted on:** 20/05/2025, **Published on:** 12/12/2025

DOI: https://doi.org/10.64091/ATICS.2025.000211

1. Introduction

In today’s information-rich digital environment, the volume of textual data published every minute, encompassing news, social media, technical reports, and more, has grown at an unprecedented pace. As a result, readers and automated systems alike find it increasingly difficult to keep up with large-scale document flows and to extract key takeaways in a concise form. To address this challenge, Shakil et al. [8] and Zhang et al. [16] have emerged as significant fields within natural language processing (NLP), aiming to condense longer documents into shorter, coherent, and fluent summaries while preserving core meaning. Broadly, ATS methods are distinguished into extractive techniques, which select important sentences or phrases verbatim from the source, and abstractive techniques, which generate new text that may paraphrase, generalise, or reorganise content to produce a summary. Abstractive summarisation is more challenging, as it requires deeper semantic understanding and the ability to produce novel output rather than simply ranking and selecting fragments of the input. In the last decade, neural sequence-to-sequence architectures (encoder–decoder) have become the predominant paradigm for abstractive summarisation. Recurrent neural networks (particularly those based on Long Short-Term Memory or LSTM units) and gated units such as GRUs were adopted early on, often augmented with attention mechanisms that enable the decoder to “focus” on relevant parts of the input at each output step. For example, attentive encoder–decoder models have shown strong performance in summarisation tasks. In today’s information-rich digital environment, the volume of textual data published every minute, encompassing news, social media, technical reports, and more, has grown at an unprecedented pace [12].

As a result, readers and automated systems alike find it increasingly difficult to keep up with large-scale document flows and to extract key takeaways in a concise form. To address this challenge, automatic text summarisation (ATS) has emerged as a significant field within natural language processing (NLP), aiming to condense longer documents into shorter, coherent, and fluent summaries while preserving core meaning. The role of pre-trained word embeddings (such as Word2Vec and GloVe) in downstream tasks, including summarisation, classification, and retrieval, has been widely analysed by Asudani et al. [3]. Embedding vectors provide rich lexical and semantic representations that help summarisation models better capture word-level meaning and improve generalisation, especially when training data is limited. Moreover, as summarisation methods mature, attention mechanisms have been shown to substantially improve performance by enabling selective encoding and decoding of relevant input segments, thereby increasing summary salience and coherence. Surveys highlight the growing importance of attention and encoder–decoder modelling for abstractive summarisation. Researchers discuss further architectural comparisons, hyper-parameter optimisation, and the integration of additional techniques to further enhance summarisation performance [22].

2. Related Work

Zhang et al. [26] developed an innovative Extract-then-Abstract framework for unsupervised abstractive summarisation. The framework is structured around a two-stage process: an initial extractive summarizer identifies key sentences, followed by a BART-based sentence-rewriting model that enhances coherence and fluency in the final abstractive summaries. By integrating both extractive and abstractive techniques, the framework aims to overcome the challenges associated with generating human-like summaries. Through a series of comprehensive experiments, including automated evaluation metrics and human assessments, the study demonstrates that the proposed framework outperforms current baselines in unsupervised summarisation. The results underscore the framework's efficacy in producing coherent, fluent summaries, thereby demonstrating its potential to advance the field of automatic document summarisation. Yousefi-Azar and Hamey [15] delve into text summarisation, specifically exploring query-oriented single-document summarisation using deep autoencoders. It investigates the impact of attention models on abstractive sentence summarisation and proposes an ensemble approach to improve the alignment between summaries and human-generated texts. The study also outlines future directions, including extending the model to generic summarisation across domains and incorporating semi-supervised learning techniques to improve summarisation performance. Mao et al. [24] focus on automatic text summarisation methods using the DUC2001 and DUC2002 datasets and evaluate them using the ROUGE metric. It highlights the superiority of these methods over supervised and unsupervised learning approaches.

Additionally, the paper emphasises the importance of prior knowledge in enhancing key-sentence selection in graph models to improve accuracy. The research paper delves into automatic text summarisation, focusing on the DUC2001 and DUC2002 datasets. It evaluates the effectiveness of these methods using the ROUGE evaluation metric, showcasing their superiority over traditional supervised and unsupervised learning techniques. Furthermore, the study underscores the importance of leveraging prior knowledge to select key sentences in graph models, thereby improving summarisation accuracy. Shang et al. [7] introduce a single, fully unsupervised framework for summarising speeches from end-to-end meetings that combines six approaches for keyword extraction, multi-sentence compression, and summarisation. The framework is designed to produce coherent summaries despite the noise in ASR transcriptions. It delivers cutting-edge functionality and can be applied to languages other than English with minimal modification. The paper also presents novel components and modules within the system, demonstrating significant advancements in meeting speech summarisation. Schumann [18] introduces an unsupervised method for sentence summarisation using a Variational Autoencoder (VAE). VAEs are recognised for learning a semantically meaningful latent variable from high-dimensional input data. By incorporating the desired output length during training, the VAE can be adjusted during inference to generate more concise summaries.

The study shows that, although these shorter summaries may not surpass a simple baseline, they yield higher ROUGE scores than reconstructing the entire sentence. The research highlights the potential of VAEs in unsupervised abstractive summarisation and suggests further investigation into improving reconstruction and the impact of length control on the model. Bražinskas et al. [1] present a novel approach to unsupervised abstractive summarisation of opinions, specifically focusing on generating summaries from collections of product reviews. Unlike traditional extractive methods that select and compile fragments from existing texts, this study introduces a generative model that creates new sentences, thereby producing more coherent and contextually relevant summaries. The authors emphasise the challenges of summarising diverse review datasets, such as Yelp and Amazon, which contain varying levels of subjective and objective information. Yang et al. [25] present TED, an unsupervised abstractive summarisation model that utilises the transformer architecture and a large-scale pretraining strategy. By leveraging lead bias in news reports and integrating theme modelling with a denoising autoencoder, TED surpasses existing unsupervised abstractive baselines across multiple benchmark datasets. The model proves effective at generating highly abstractive summaries without supervised training on labelled data. Nguyen et al. [23] propose an unsupervised method for generating class-specific abstractive summaries of customer reviews, tackling the challenge of efficiently summarising large volumes of online reviews without relying on human-written summaries.

The proposed model, Class-Copy Cat, uses a generative variational autoencoder with a hierarchical structure and a class-correlation gating mechanism to capture dependencies among products, reviews, and classes. This approach allows for the generation of summaries that are highly relevant, fluent, and representative of specific product aspects. Human evaluation and comparisons with reference datasets demonstrate the model's superiority over existing abstractive and extractive summarisation baselines, representing a major advancement in unsupervised opinion summarisation. Zhuang et al. [9] present SCR (Summarise, Contrast, and Review), an unsupervised learning technique for abstractive summarisation. Unlike traditional methods that rely on large-scale parallel corpora and supervised learning, SCR leverages contrastive learning to generate summaries without requiring reference summaries. The method includes a summarizer, a contrastive encoder to ensure semantic closeness to source documents, and a writing reviewer to maintain summary quality. Experimental results demonstrate the effectiveness of SCR, and future work aims to explore transferable feature learning and semi-supervised learning. Challagundla and Peddavenkatagari [2] present a neural sequence-to-sequence model aimed at improving abstractive text synopsis. The model leverages advanced deep learning techniques to effectively handle rare and out-of-vocabulary (OOV) words while maintaining the coherence and informativeness of the generated summaries. Through rigorous training and comparative analysis, the model demonstrates significant performance improvements over existing state-of-the-art methods [19].

Abstractive summarisation is an important field of study in Natural Language Processing (NLP) that aims to generate concise, coherent, and information-rich summaries. Gupta and Gupta [21] provide an overview of recent studies and advancements in abstractive summarisation, highlighting techniques, tools, evaluation measures, challenges, and future trends in the field. By categorising works and discussing the popular components of abstractive summarisation systems, this study aims to systematise knowledge and enhance understanding in this research domain. The paper addresses the complexities of natural language text that make abstractive summarisation challenging, emphasising the importance of producing high-quality, grammatically correct, and readable summaries. Shaker et al. [6] focus on text summarisation strategies, comparing extractive and abstractive methods using various deep learning techniques, including Fuzzy Logic, LSTM-CNN, and transformers. It discusses the challenges and benefits of summarising large texts, emphasising the importance of retaining essential information. The study evaluates different methodologies, presents findings from 16 research publications, and offers recommendations to enhance text summarisation techniques. Gehrmann et al. [20] propose abstractive summarisation-based techniques for neural networks that aim to improve fluency but struggle with content selection. This study introduces a bottom-up approach using a data-efficient content selector to enhance text compression and fluency in summaries. By combining this approach with abstractive models, the research demonstrates significant improvements in ROUGE scores on the CNN-DM and NYT datasets, highlighting the effectiveness of the two-step process for generating concise, fluent summaries.

Coavoux et al. [11] focus on unsupervised opinion summarisation using an abstractive neural summarisation system. The system includes a language model to encode reviews and generate sentences, as well as a clustering step to group reviews by aspects for focused summary generation. Experiments on the Opossum dataset demonstrate the importance of clustering and leveraging aspect information for improved summarisation. Paulus [17] introduces a novel neural network model for abstractive summarisation that addresses the issue of repetitive or illogical phrases in longer documents. By integrating supervised word prediction with reinforcement learning, the model reduces "exposure bias" and improves the readability of the generated summaries. The model achieves results on datasets such as CNN/Daily Mail and the New York Times, offering a new attention mechanism and learning objective to enhance summarisation quality. Socher Saiyyad and Patil [13] provide a comprehensive overview of the advancements in text summarisation methodologies, particularly focusing on deep learning approaches. It discusses the limitations of traditional summarisation techniques, which often struggle with semantic understanding and context retention. The authors highlight various deep learning models, including neural networks and hybrid systems, that have been developed to enhance the accuracy and effectiveness of summarisation across multiple languages.

The review also examines the application of various algorithms, such as restricted Boltzmann machines and K-nearest neighbours, for generating concise and coherent summaries. By synthesising findings from various studies, the paper underscores the potential of deep learning to revolutionise text summarisation, offering insights into future research directions and into integrating advanced techniques to address existing challenges. The authors conclude that while significant progress has been made, further exploration is needed to refine these models and improve their applicability in real-world scenarios. Chu and Liu [4] proposed a novel approach to unsupervised multi-document abstractive summarisation, specifically targeting review data from platforms such as Yelp and Amazon. The authors propose a neural model that generates coherent, relevant summaries without requiring extensive document-summary pairs, which are often costly to obtain. Through a series of ablation studies, the importance of various architectural components is demonstrated, highlighting the model's ability to produce highly abstractive, fluent, and sentiment-representative summaries. The evaluation process includes both automated metrics, such as ROUGE scores, and human assessments, confirming the model's superiority over traditional extractive baselines. The findings suggest that the proposed model effectively addresses the limitations of existing supervised methods, paving the way for future research in unsupervised summarisation techniques. The paper concludes with a discussion of potential improvements and applications of the model across different summarisation contexts. Nayeem et al. [14] developed an unsupervised abstractive summarisation system for multi-document settings.

The authors propose a paraphrastic sentence fusion model that combines sentence fusion and paraphrasing using skip-gram word embeddings. The model aims to improve the information coverage and abstractiveness of the generated sentences. Experimental evaluations are conducted on human-generated multi-sentence compression datasets and on datasets from different domains for multi-document summarisation. The results show significant improvements over state-of-the-art methods. The paper also addresses the problem of the classical summary-length limit in multi-document summarisation. The related work section provides an overview of abstractive summarisation techniques, including compressive summarisation, sentence fusion, and neural network-based approaches. The authors highlight the limitations of existing models and emphasise the need for unsupervised, full abstractive models. The paper introduces novel techniques for sentence fusion and paraphrasing based on skip-gram word embeddings. The proposed model is evaluated on multiple datasets and shown to achieve state-of-the-art results. The paper concludes with a discussion of future research directions. Supriyono et al. [22] delve into advancements in summarisation models, focusing on the effectiveness of various techniques for generating accurate and informative summaries. It explores the impact of different decoding methods, such as beam-search and nucleus sampling, on the quality of generated text. The study compares the performance of different models, including an autoencoder model that tends to fully copy input sentences, affecting the ability to derive meaningful latent codes. The research highlights the drawbacks of beam-search decoding, which may lead to repetitive or bland text generation, compared with nucleus sampling, which enhances formativeness and improves ROUGE-1 scores while reducing self-BLEU.

Evaluation criteria include coverage of common opinions, alignment with input reviews, and the ability to generate informative content across diverse topics. An ablation study is conducted to analyse the impact of individual components on summarisation performance, including the presence of a discriminator, an attention mechanism, and nucleus sampling. Overall, the research paper explores the nuances of summarisation models, decoding strategies, model comparisons, and evaluation metrics to enhance the efficiency and effectiveness of automated text summarisation. TASP for topic detection and abstractive summarisation of social media posts [10]. It leverages transformer-based architectures to identify clusters of social posts on similar topics and generate concise summaries that allow for human exploration of topic-specific information. The study highlights the limitations of current summarisation approaches, the inability to leverage visual cues and multi-modal content, and suggests future research directions to address these challenges. The paper introduces the concept of abstractive multi-post summarisation and demonstrates the effectiveness of the PRIMERA summarizer in handling large collections of posts. One limitation of the methodology is its assumption that each post covers at most one topic, which may introduce some noise into the input data. Future research is recommended to address multi-topic, multi-post summarisation for more accurate representations. Semantic Abstractive Summarisation (SAS). SAS, as introduced by Liu et al. [5]. First, generate an Abstract Meaning Representation (AMR) graph of an input story, through which they extract a summary graph, and finally, create summary sentences from this summary graph [19].

3. Methodology

The overall block diagram for the Abstractive Text Summarisation model is shown in Figure 1. To provide Abstractive Text Summarization for news articles through the use of an LSTM-based encoder and Decoder structure. For training purposes, the InShorts News Article Dataset has been identified. This dataset contains headlines and summaries of news items, along with their sources.

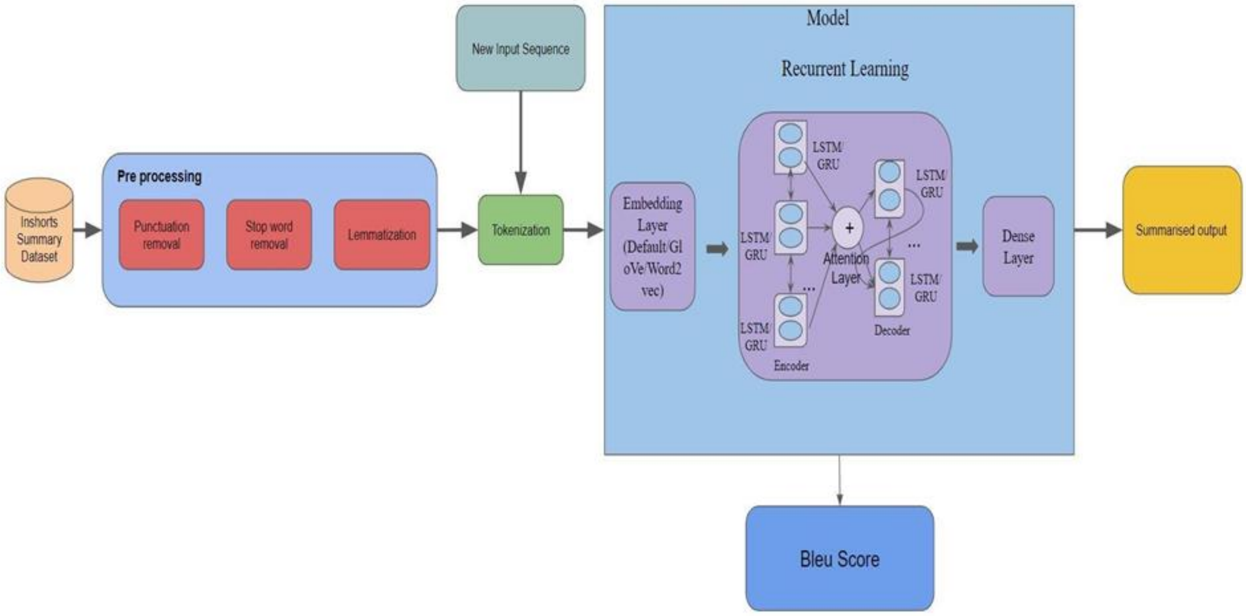

Figure 1: Block diagram for abstractive text summarisation model

This dataset contains 55104 articles and their headlines for training:

- **Word Embeddings GloVe:** GloVe is an unsupervised learning algorithm for obtaining vector representations for words. Training is performed on aggregated global word-word co-occurrence statistics from a corpus, and the resulting representations showcase interesting linear substructures of the word vector space.
- **Word2Vec:** Word2Vec is not a single algorithm but a collection of model architectures and optimisation techniques for learning word embeddings from large datasets. These embeddings have demonstrated their effectiveness across a wide range of downstream natural language processing tasks.
- **LSTM:** Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) architecture widely used in deep learning. Unlike traditional feed-forward neural networks, LSTMs include feedback connections, enabling them to handle not only individual data points (e.g., images) but also sequences of data (e.g., speech or video). An LSTM unit consists of a cell, an input gate, an output gate, and a forget gate. The cell serves as a memory element, retaining information over extended time intervals, while the gates control the flow of information entering, leaving, or being retained within the cell.
- **Gated Recurrent Units GRU:** Gated Recurrent Units (GRUs) are a gating mechanism used in recurrent neural networks. Similar to Long Short-Term Memory (LSTM) networks, GRUs include a forget gate but are more streamlined, as they lack an output gate, resulting in fewer parameters. Despite this simplification, GRUs have demonstrated performance comparable to LSTMs across tasks such as polyphonic music modelling, speech signal modelling, and natural language processing.
- **Attention Layer:** An Attention Layer enables a model to analyse the entire information contained in the original sentence and generate the most appropriate word based on the current word and its context. It also allows the model to adjust its focus, zooming in on specific details or zooming out to consider broader features, thereby capturing both local and global patterns.

3.1. Vanilla BiGRU

The BiGRU-based autoencoder-decoder model shown in Figure 2 is designed to take news articles as input and generate their corresponding summaries. Before feeding the data into the model, it was cleaned using preprocessing strategies such as stop-word removal and lemmatisation. Word embeddings are learned during model training.

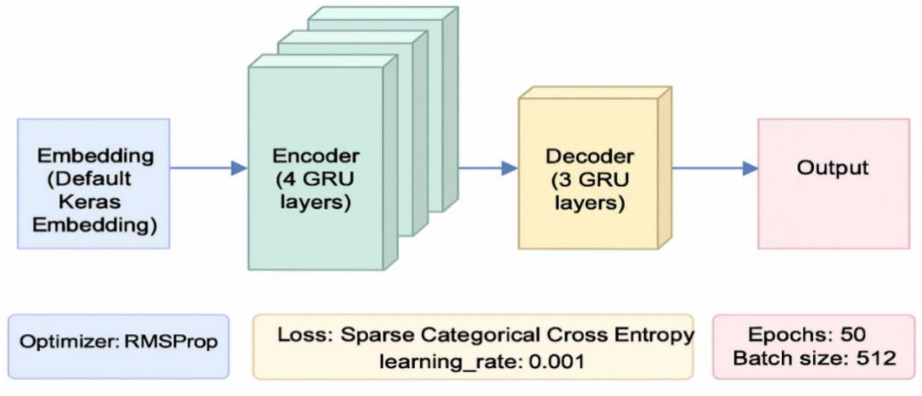

Figure 2: A BiGRU based auto encoder - decoder model

3.2. Vanilla LSTM

The LSTM-based autoencoder-decoder model shown in Figure 3 is designed to take a news article as input and generate its corresponding summary. Before feeding the data into the model, it was cleaned using preprocessing strategies such as stop-word removal and lemmatisation. Word embeddings are learned during model training.

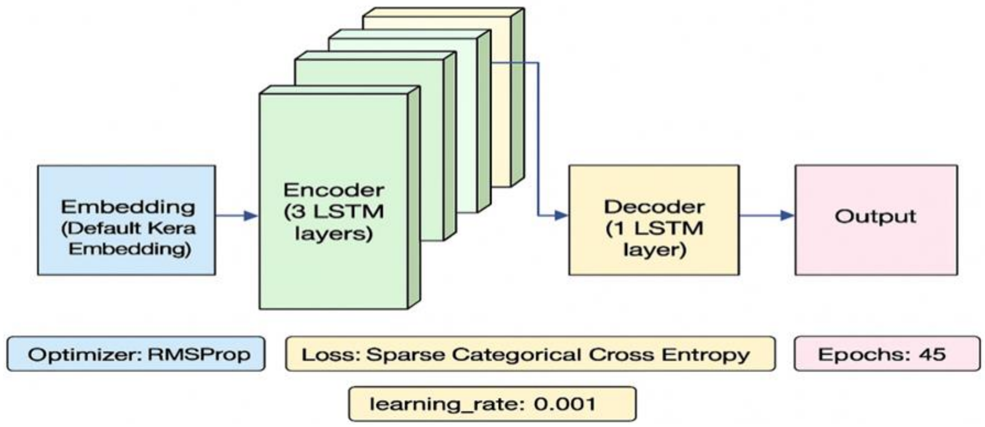

Figure 3: ALSTM based auto encoder - decoder model

3.3. LSTM with Glove Embedding

A LSTM-based autoencoder-decoder model with GloVe embeddings, as in Figure 4, has been designed to take a news article as input and generate its corresponding summary. A pre-trained 300-dimensional GloVe model is used to generate an embedding matrix, which is then used during training of both the encoder and decoder. Before feeding the data into the model, it was cleaned using preprocessing strategies such as stop-word removal and lemmatisation.

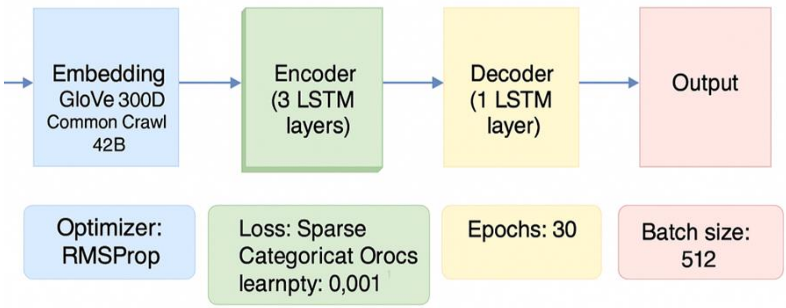

Figure 4: An LSTM–based autoencoder–decoder model with GloVe embedding

3.4. LSTM with Attention and Glove Embedding

A LSTM-based autoencoder-decoder model, shown in Figure 5, with attention and GloVe embeddings, has been designed to take a news article as input and generate its corresponding summary. A pre-trained 300-dimensional Glove model is used to generate an embedding matrix, which is then used during training of both the encoder and decoder. Additionally, an attention layer is added between the encoder and decoder outputs so the model knows which parts of the encoder outputs to pay more attention to. Before feeding the data into the model, it was cleaned using preprocessing strategies such as stop-word removal and lemmatisation.

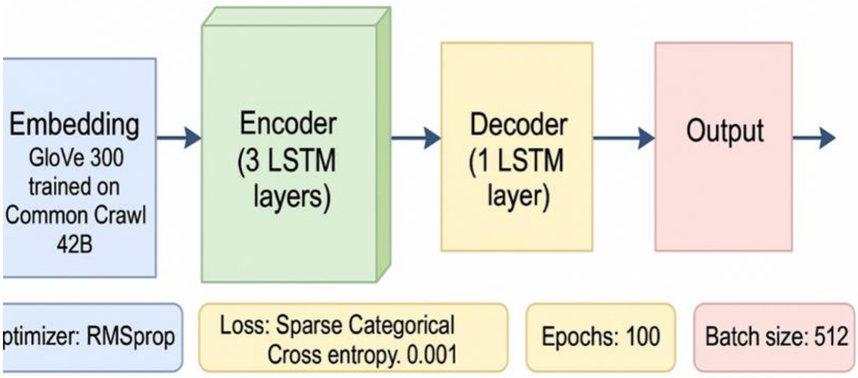

Figure 5: An LSTM-based autoencoder-decoder model with attention and GloVe embedding

3.5. LSTM with Attention and Word2vec Embedding

A LSTM-based autoencoder-decoder model, as in Figure 6, with attention and word2vec embeddings, has been designed to take a news article as input and generate its corresponding summary. Pretrained Word2vec embeddings from the gensim module are used to generate the embedding matrix, which is then used during training of both the encoder and decoder. Additionally, an attention layer is added between the encoder and decoder outputs so the model knows which parts of the encoder outputs to pay more attention to. Before feeding the data into the model, it was cleaned using preprocessing strategies such as stop-word removal and lemmatisation.

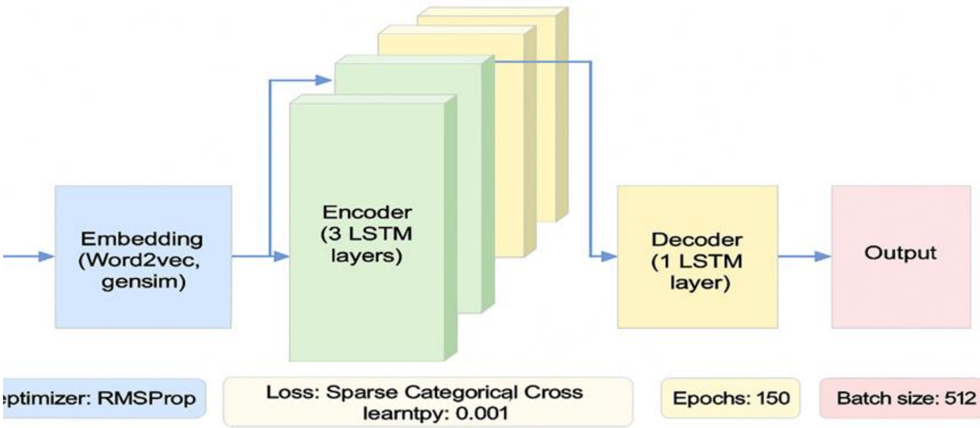

Figure 6: A LSTM based auto encoder - decoder model with attention and word2vec embedding

3.6. Vanilla GRU

An AGRU-based autoencoder-decoder model, shown in Figure 7, has been designed to take a news article as input and generate its corresponding summary. Before feeding the data into the model, it was cleaned using preprocessing strategies such as stop-word removal and lemmatisation. Word embeddings are learned during model training.

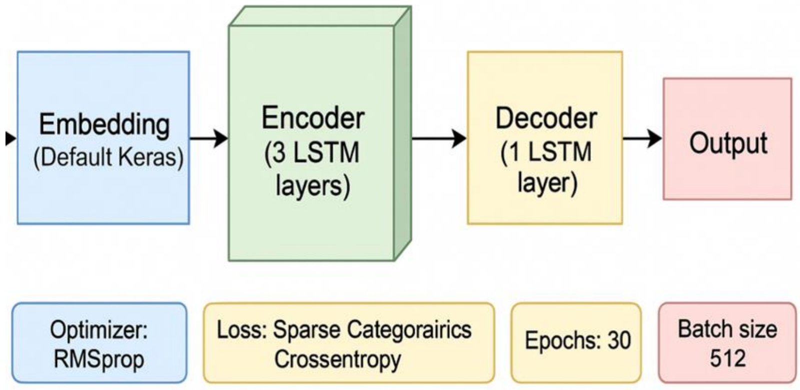

Figure 7: AGRU based auto encoder - decoder model

3.7. GRU with GloVe Embedding

A GRU-based autoencoder-decoder model in Figure 8, using GloVe embeddings, has been designed to take a news article as input and generate its corresponding summary. A pre-trained 300-dimensional Glove model is used to generate an embedding matrix, which is then used during training of both the encoder and decoder. Before feeding the data into the model, it was cleaned using preprocessing strategies such as stop-word removal and lemmatisation.

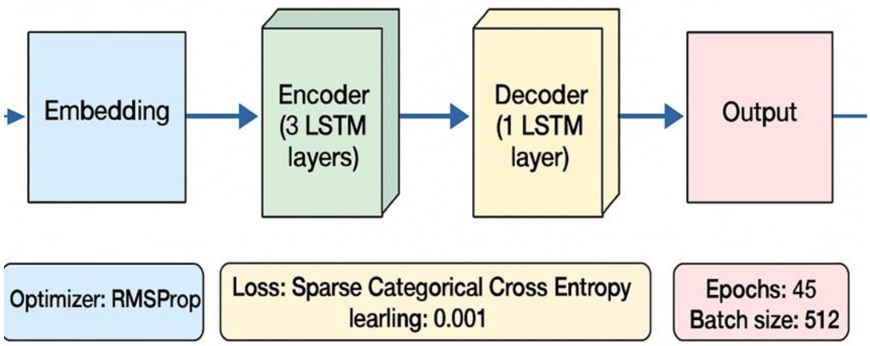

Figure 8: A GRU-based autoencoder-decoder model with GloVe embedding

3.8. GRU with Word2vec Embedding

A GRU-based autoencoder-decoder model in Figure 9 with Word2Vec embeddings has been designed to take a news article as input and generate its corresponding summary.

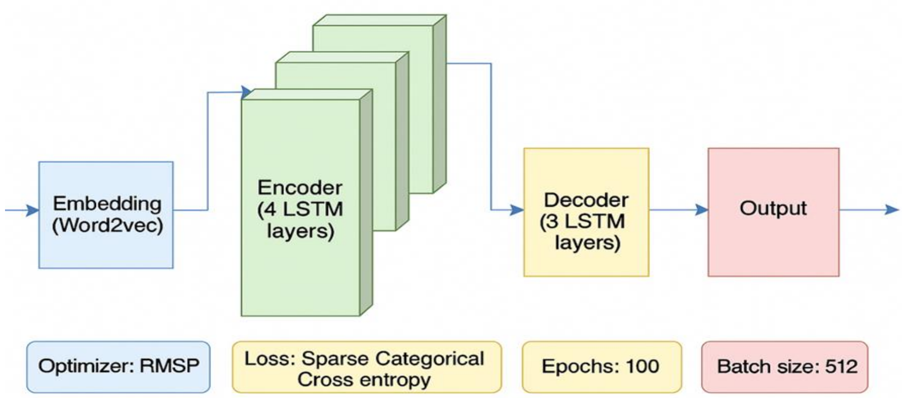

Figure 9: AGRU-based autoencoder-decoder model with word2vec embedding

Pretrained Word2vec embeddings from the gensim module are used to generate the embedding matrix, which is then used during training of both the encoder and decoder. Before feeding the data into the model, it was cleaned using preprocessing strategies such as stop-word removal and lemmatisation.

4. Results and Discussion

The BLEU score is a performance metric for evaluating text generation. It compares a candidate translation to one or more reference translations. While originally designed for machine translation, BLEU is also applicable to assessing text generated for various natural language processing tasks. Ideally, the Perfect BLEU Score is given by 1:

$$\mathrm{Bleu} = \rho . e^{\sum_{n=1}^{N} \frac{1}{N} * (\log \rho\ n)}$$

ρ=Brevity Penalty given by:$(1 - c) - \ \rho = 1 \text{ if } c > r \text{ else } e$

Table 1: Bleu score of each model

Model	Bleu Score
Model 1: BiGRU Vanilla (Paper Implementation)	0.5748
Model 2: LSTM Vanilla	0.5504
Model 3: LSTM with GloVe	0.6016
Model 4: LSTM with GloVe and Attention	0.7254
Model 5: LSTMwithWord2VecandAttention	0.7481
Model 6: GRU Vanilla	0.5320
Model 7: GRU with GloVe	0.6259
Model 8: GRUwithWord2Vec	0.6208

The Bleu scores for each model have been recorded and tabulated in Table 1 below, and the model Bleu scores are shown in Figure 10.

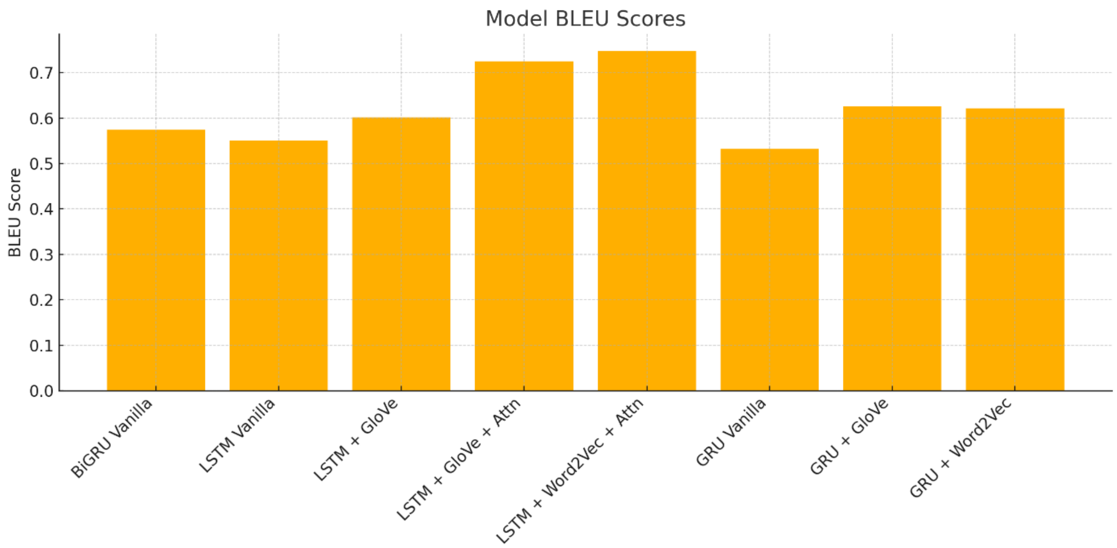

Figure 10: Model bleu scores

As shown in Table 1, Model5, which uses an LSTM-based Encoder-Decoder Architecture with an Attention Layer, achieves the best BLEU score of 0.7481. BLEU scores measure the performance of various models on a task, typically related to text generation or translation. Here is a synthesis of insights and relevant information based on these scores and the context from recent findings on similar models:

- BLEU scores range from about 0.53 to 0.75, with the highest scores belonging to LSTM-based models using Word2Vec embeddings combined with Attention (0.7481), followed by LSTM with GloVe and Attention (0.7254).

Vanilla models (basic LSTM, BiGRU, GRU without embeddings or attention) score substantially lower, between 0.53 and 0.57 for BiGRU and GRU, and 0.55 for LSTM vanilla.

- Incorporating pretrained embeddings such as GloVe or Word2Vec significantly improves BLEU scores for LSTM and GRU models. The use of Attention mechanisms further improves the LSTM model's performance. It aligns with the literature, which shows that LSTM models with pretrained embeddings and attention mechanisms consistently outperform vanilla RNNs, GRUs, or BiGRUs without such enhancements. Therefore, Performance gains in GRU variants with embeddings are present but slightly lower than those in LSTM counterparts with embeddings and attention. The BiGRU Vanilla (paper implementation) score of 0.5748 generally aligns with typical vanilla BiGRU model performance reported in classification or sequence tasks. However, some studies show that BiGRU models outperform vanilla LSTMs or GRUs, but usually without attention or embeddings.
- The BLEU scores illustrate a common trend: vanilla recurrent models (LSTM, GRU, BiGRU) have moderate baseline performance; employing pretrained word embeddings like GloVe or Word2Vec substantially improves performance; and adding attention mechanisms to LSTM models provides the greatest improvement in BLEU score among the variants tested. These observations align well with current deep learning research on sequence modelling and natural language processing.

For this model, some example summaries are provided in Figure 11.

```
print("Review:",seq2text(padded_xtest[13]))
print("Original summary:",seq2summary(padded_ytest[13]))
print("Predicted summary:",decode_sequence(padded_xtest[13]))
```

```
Review: rafael benitez appoint new manager relegation threaten premier league outfit newcastle united magpies sack previous manager steve mclaren early today benitez sack real madrid january seven month charge sign three year deal newcastle previously manage liverpool inter milan napoli valencia
Original summary: rafael benitez hired as newcastle united manager
Predicted summary:  rafael benitez hired as newcastle united manager
```

```
print("Review:",seq2text(padded_xtest[25]))
print("Original summary:",seq2summary(padded_ytest[25]))
print("Predicted summary:",decode_sequence(padded_xtest[25]))
```

```
Review: mall china 39 s shanxi create rooster statue resemble us president elect donald trump small replicas statue available sale price start 3 800 11 lakh 32 foot tall statue notably new year china begin january 28 year rooster accord chinese zodiac
Original summary: china mall installs 39 donald trump like 39 rooster statue
Predicted summary:  china mall installs 39 donald trump statue 39 china
```

```
print("Review:",seq2text(padded_xtest[48]))
print("Original summary:",seq2summary(padded_ytest[48]))
print("Predicted summary:",decode_sequence(padded_xtest[48]))
```

```
Review: usa become first nation register gold medal rio olympics 2016 19 year old virginia thrasher gold 10 metre air rifle day one event hold saturday thrasher finish ahead china 39 s du li yi sile silver bronze medal open china 39 s medal tally event
Original summary: usa wins first gold medal of rio olympics 2016
Predicted summary:  usa wins first gold medal at rio olympics
```

Figure 11: Summaries of the models

5. Conclusion and Future Work

This study demonstrates the effectiveness of using Long Short-Term Memory (LSTM) networks with an encoder-decoder architecture for abstractive text summarisation of news articles. The integration of pre-trained word embeddings (Word2Vec) and attention mechanisms significantly improves the quality of the generated summaries, with the highest BLEU score (0.7481) achieved using an LSTM encoder-decoder model incorporating both features. Our results highlighted the importance of attention layers and pre-trained embeddings in enhancing the semantic coherence and fluency of summaries. Additionally, we find that both LSTMs and GRUs, despite their differences, are effective at handling sequential data for summarisation tasks. This research underscores the potential of automated summarisation techniques for more efficient information retrieval and better content comprehension. Several promising directions for future research in abstractive text summarisation are identified. First, exploring architectural comparisons across recurrent and transformer-based models, such as the Transformer architecture, could offer insights into how newer models perform relative to traditional LSTM-based approaches. A more detailed analysis of hyperparameters, such as learning rates, batch sizes, and sequence lengths, may reveal opportunities for further model improvements. Additionally, using more advanced pre-trained embeddings, such as BERT or GPT models, could further enhance summary quality. Incorporating techniques like data augmentation or semi-supervised learning could improve model robustness, particularly when working with diverse datasets. Beyond automated metrics like BLEU scores, human evaluation metrics—such as readability, relevance, and coherence—could provide a more comprehensive assessment of summary quality. Expanding this research to multilingual or cross-lingual summarisation could unlock new opportunities to improve summarisation systems globally. By addressing these areas, future studies can push the boundaries of abstractive text summarisation, leading to more accurate, readable, and useful automated summaries.

Acknowledgement: The authors express their gratitude to Saveetha Engineering College, SRM Institute of Science and Technology, KCG College of Technology, Francis Xavier Engineering College, and the University of Texas at Dallas for their academic guidance and institutional support.

Data Availability Statement: The datasets supporting the conclusions of this work are available from the corresponding authors upon reasonable request, subject to applicable data-sharing policies.

Funding Statement: This research received no financial support.

Conflicts of Interest Statement: The authors declare no conflicts of interest that could potentially influence the research findings.

Ethics and Consent Statement: All contributing authors have reviewed and approved the final manuscript and consent to its publication for research and educational purposes.

Reference

1. A. Bražinskas, M. Lapata, and I. Titov, “Unsupervised opinion summarization as copycat-review generation,” *arXiv preprint*, 2019. Available: https://arxiv.org/abs/1911.02247 [Accessed by 11/10/2024].
2. B. C. Challagundla and C. Peddavenkatagari, “Neural sequence-to-sequence modelling with attention by leveraging deep learning architectures for enhanced contextual understanding in abstractive text summarization,” *arXiv preprint*, 2024. Available: https:// arxiv.org/abs/2404.08685 [Accessed by 30/10/2024].
3. D. S. Asudani, N. K. Nagwani, and P. Singh, “Impact of word embedding models on text analytics in deep learning environment: A review,” *Artificial Intelligence Review*, vol. 56, no. 2, pp. 10345–10425, 2023.
4. E. Chu and P. Liu, “MeanSum: A neural model for unsupervised multi-document abstractive summarization,” in *Proc. Int. Conf. Machine Learning (ICML)*, Long Beach, California, United States of America, 2019.
5. F. Liu, J. Flanigan, S. Thomson, N. Sadeh, and N. A. Smith, “Toward abstractive summarization using semantic representations,” *arXiv preprint*, 2018. Available: https://arxiv.org/abs/1805.10399 [Accessed by 12/10/2024].
6. F. Shaker, M. M. Abd Zaid, A. A. T. Al-Khazaali, and Z. A. Al-Khzaali, “Text summarization using different methods for deep learning,” *BIO Web of Conferences*, vol. 97, no. 4, p. 00074, 2024.
7. G. Shang, W. Ding, Z. Zhang, A. Tixier, P. Meladianos, M. Vazirgiannis, and J. P. Lorré, “Unsupervised abstractive meeting summarization with multi-sentence compression and budgeted submodular maximization,” *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Melbourne, Australia, 2018.
8. H. Shakil, A. Farooq, and J. Kalita, “Abstractive text summarization: State of the art, challenges, and improvements,” *Neurocomputing*, vol. 603, no. 10, p. 128255, 2024.
9. H. Zhuang, W. E. Zhang, J. Yang, C. Ma, Y. Qu, and Q. Z. Sheng, “Learning from the source document: Unsupervised abstractive summarization,” in *Findings of the Assoc. for Computational Linguistics: EMNLP*, Abu Dhabi, United Arab Emirates, 2022.
10. I. Benedetto, M. La Quatra, L. Cagliero, L. Vassio, and M. Trevisan, “TASP: Topic-based abstractive summarization of Facebook text posts,” *Expert Systems with Applications*, vol. 255, no. 12, p. 124567, 2024.
11. M. Coavoux, H. Elsahar, and M. Gallé, “Unsupervised aspect-based multi-document abstractive summarization,” in *Proc. 2nd Workshop on New Frontiers in Summarization*, Hong Kong, China, 2019.
12. M. Isonuma, J. Mori, D. Bollegala, and I. Sakata, “Unsupervised abstractive opinion summarization by generating sentences with tree-structured topic guidance,” *Trans. Assoc. Comput. Linguistics*, vol. 9, no. 9, pp. 945–961, 2021.
13. M. M. Saiyyad and N. N. Patil, “Text summarization using deep learning techniques: A review,” *Engineering Proceedings*, vol. 59, no. 1, p. 194, 2024.
14. M. T. Nayeem, T. A. Fuad, and Y. Chali, “Abstractive unsupervised multi-document summarization using paraphrastic sentence fusion,” in *Proceedings of the 27th International Conference on Computational Linguistics*, Santa Fe, New Mexico, United States of America, 2018.
15. M. Yousefi-Azar and L. Hamey, “Text summarization using unsupervised deep learning,” *Expert Systems with Applications*, vol. 68, no. 2, pp. 93–105, 2017.

16. M. Zhang, G. Zhou, W. Yu, N. Huang, and W. Liu, "A comprehensive survey of abstractive text summarization based on deep learning," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, p. 7132226, 2022.
17. R. Paulus, "A deep reinforced model for abstractive summarization," *arXiv preprint*, 2017. Available: https://arxiv. org/abs/1705. 04304 [Accessed by 15/10/2024].
18. R. Schumann, "Unsupervised abstractive sentence summarization using length-controlled variational autoencoder," *arXiv preprint*, 2018. Available: https://arxiv.org/abs/1809.05233 [Accessed by 20/10/2024].
19. S. Dohare, V. Gupta, and H. Karnick, "Unsupervised semantic abstractive summarization," in *Proc. ACL Student Research Workshop*, Melbourne, Australia, 2018.
20. S. Gehrmann, Y. Deng, and A. M. Rush, "Bottom-up abstractive summarization," *arXiv preprint*, 2018. Available: https://arxiv. org/abs/1808.10792 [Accessed by 17/10/2024].
21. S. Gupta and S. K. Gupta, "Abstractive summarization: An overview of the state of the art," *Expert Systems with Applications*, vol. 121, no. 5, pp. 49–65, 2019.
22. S. Supriyono, A. P. Wibawa, and F. Kurniawan, "A survey of text summarization: Techniques, evaluation and challenges," *Natural Language Processing Journal*, vol. 7, no. 6, p. 100070, 2024.
23. T. N. A. Nguyen, M. Shen, and K. Hovsepian, "Unsupervised class-specific abstractive summarization of customer reviews," in *Proc. 4th Workshop on e-Commerce and NLP*, Bangkok, Thailand, 2021.
24. X. Mao, H. Yang, S. Huang, Y. Liu, and R. Li, "Extractive summarization using supervised and unsupervised learning," *Expert Systems with Applications*, vol. 133, no. 11, pp. 173–181, 2019.
25. Z. Yang, C. Zhu, R. Gmyr, M. Zeng, X. Huang, and E. Darve, "TED: A pretrained unsupervised summarization model with theme modelling and denoising," *arXiv preprint*, 2020. Available: https://arxiv.org/abs/2001.00725 [Accessed by 25/10/2024].
26. Z. Zhang, X. Liang, Y. Zuo, and Z. Li, "Unsupervised abstractive summarization via sentence rewriting," *Computer Speech & Language*, vol. 78, no. 3, p. 101467, 2023.