**SummaryFlow: AI Book Summarization**

**Ahmed Hany**  
Scientific Computing Department,  
Ain Shams UniversityCairo, Egypt  
20201700090@cis.asu.edu.eg

**Ahmed Wail**  
Scientific Computing Department,  
Ain Shams UniversityCairo, Egypt  
20201700091@cis.asu.edu.eg **George Saeid**  
Scientific Computing Department,  
Ain Shams UniversityCairo, Egypt  
20201700210@cis.asu.edu.eg

**Abdelrahman Mohamed**Scientific Computing Department,  
Ain Shams UniversityCairo, Egypt  
20201700483@cis.asu.edu.eg **Esraa Saad**Scientific Computing Department,  
Ain Shams UniversityCairo, Egypt  
20201701055@cis.asu.edu.eg

**Ganna Ayman**Scientific Computing Department,  
Ain Shams UniversityCairo, Egypt  
20201700204@cis.asu.edu.eg

Abstract:

Long narrative novel summarization involves creating a concise and meaningful summary of a novel by including relevant and important points. There are two approaches to this: Extractive Summarization, which involves selecting the most relevant statements from the novel, and Abstractive Summarization, which involves generating new sentences for the summary. A significant challenge in scaling machine learning is training models to perform tasks that are difficult or time-consuming for humans to evaluate. We address this challenge in the context of abstractive summarization of long narrative novels. Our method uses recursive task decomposition, where the model first summarizes small sections of the novel and then recursively summarizes these summaries to produce a summary of the entire novel. We have fine-tuned the FLAN-T5 base model on the BookSum dataset for summarizing book-length narratives.

Keywords—Summarization, Extractive, Abstractive, Transformer

# **Introduction**

With the advent of digital reading devices and online platforms, accessing novels has become easier and more widespread among the public. However, the sheer volume of available books can make finding relevant or interesting novels a daunting task. Readers often experience information overload when sifting through countless book summaries and reviews. To mitigate this, automatic novel summarization has become a valuable tool, quickly highlighting the core themes and key points of a book. Novel summarization can be divided into two main approaches: extractive and abstractive. Extractive summarization involves selecting important sentences directly from the book, while abstractive summarization requires a deeper understanding and rephrasing of the content, presenting a unique challenge in capturing the essence of a novel.

Deep learning models show promise in addressing the challenges of novel book summarization. Earlier models like Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) struggled with issues such as vanishing gradients and slower training times. However, the advent of Transformer-based architectures has been a breakthrough, significantly enhancing the techniques used for summarizing novels. These advanced models are better equipped to understand and condense complex narratives, providing more accurate and coherent summaries.

Abstractive book summarization is particularly difficult due to the challenges in dataset collection ‎[1], Existing methods are typically either extractive ‎[2] or focus on shorter stories ‎[3].

We implement a natural task decomposition for long-form novel summarization: first, we train models to summarize small sections of a book, then use these models to summarize larger sections, continuing this process recursively. To achieve this, we utilize the T5 transformer architecture, known for its flexibility and effectiveness in text generation tasks. We train a single model for these tasks using standard cross-entropy behavioral cloning (BC), allowing the T5 model to efficiently handle the recursive summarization process and produce coherent summaries for entire novels.

# **Related Work**

Most of the previous research in the field of text summarization has primarily focused on short text summarization. Techniques for summarizing articles, news reports, and other brief documents have been extensively developed and refined. These studies often leverage extractive and abstractive methods to generate concise summaries that capture the essential information from short texts. However, the challenge of long text summarization, particularly for novels and other lengthy narrative forms, has not been as thoroughly addressed.

The complexity of summarizing long texts arises from several factors. Firstly, maintaining coherence and cohesiveness over a longer narrative is significantly more challenging than with shorter texts. Long texts often involve multiple subplots, complex character development, and intricate thematic elements that need to be accurately represented in the summary. Secondly, the sheer volume of content means that summarization methods must effectively condense vast amounts of information without losing the essence of the story.

One of the approaches which we have is the one done by Porwal et al. 2022 ‎[4]*,* The article discusses the use of Transformer-based models, particularly BERT, for automatic book summarization, emphasizing their superiority over traditional machine learning approaches. It differentiates between extractive and abstractive summarization methods, focusing on the challenges and advancements in abstractive techniques. The paper also presents an implementation of BERT ‎[5] for chapter-wise extractive summarization, ensuring the preservation of the book's context. To evaluate the performance of the model-generated summaries against human-generated abstracts, it utilizes ROUGE ‎[6] scores, aiming to enhance text generation to match human cognitive levels.

Other approach which has made by Wu et al. (2020) ‎[7], They present a novel approach to summarizing long narrative texts using human feedback to enhance machine learning models. The core methodology builds on previous reinforcement learning techniques, specifically task decomposition and recursive reward modeling. By breaking down the summarization task into smaller, manageable sub-tasks and iterating with human feedback at each level, the authors demonstrate a significant improvement in summarization quality for complex and lengthy documents such as books.

The authors acknowledge the foundational work in applying human feedback to reinforcement learning, citing Christiano et al. (2017, 2018) ‎[8] and Leike et al. (2018) ‎[9] as key influences. They also draw parallels with other domains where human feedback has been effectively integrated, including dialogue systems, translation, and story generation.

A distinctive aspect of their approach is the application to the NarrativeQA dataset, which consists of question-answer pairs about full books and movie transcripts. The research shows that their recursive summarization method produces summaries that can be effectively used for question answering, achieving competitive results even when compared to models explicitly trained for this task. This indicates that their summaries retain a high level of detail and accuracy, essential for practical applications like information retrieval and content understanding.

# **Approach**

In tackling the complex problem of summarizing long narrative novels, our approach leverages and extends state-of-the-art techniques to achieve high-quality, concise summaries. Given the intricate nature of narrative structures and the need to preserve critical plot points and character development, a straightforward summarization approach often falls short. To address these challenges, we have devised a comprehensive methodology that incorporates recursive task decomposition and advanced preprocessing techniques. This approach not only builds on the foundational work of prior research but also introduces a novel pipeline tailored to the unique demands of summarizing extensive literary works. Our method ensures that the final summaries are both informative and manageable, maintaining the essence of the original text while reducing it to a more accessible form.

**3.1: Task Decomposition**

In our approach, we have adopted and extended the methodology presented by Wu et al. (2020) ‎[7]. The foundational principle of task decomposition, as employed by Wu et al., involves breaking down the complex task of summarizing a long narrative into smaller, manageable sub-tasks. This recursive process allows for more effective handling of extensive texts by summarizing smaller sections incrementally and then combining these summaries into a comprehensive overview.

## Building on this, we have designed a new pipeline that enhances and adapts the original task decomposition technique to better suit our specific requirements for summarizing long narrative novels. Our modifications aim to streamline the process and improve the efficiency and accuracy of the summarizations. Units

## **3.2: Recursive Decomposition for Book Summarization**

**3.2.1: PreProcessing**

**Objective:**

The primary goal of preprocessing is to transform raw PDF documents into a format suitable for text summarization. This involves extracting and cleaning the text to ensure that the input to the summarization model is of high quality and free from extraneous content.

**Steps:**

**Text Extraction:**

The text content is extracted from the PDF document. This involves converting the PDF format into a text format that can be manipulated and analyzed.

Noise Removal:

Unwanted elements such as tables of contents, indices, footnotes, and headers/footers are identified and removed. This is crucial to eliminate any non-relevant information that could degrade the performance of the summarization model.

The text is normalized by converting it to a consistent format. This includes standardizing fonts, removing special characters, correcting typos, and ensuring consistent punctuation. Normalization helps in reducing variability in the text, which can improve the model's performance.

Tokenization:

The normalized text is tokenized into individual words or sub words. Tokenization is a fundamental step for natural language processing tasks as it converts the text into a format that the summarization model can process effectively.

**Outcome:**

The preprocessing stage produces a clean, segmented, and tokenized text that serves as high-quality input for the summarization model. This ensures that the model focuses on the relevant content and generates more accurate and coherent summaries.

**Segmentation:**

Once preprocessing is complete, the text is passed to a fixed chunking algorithm. This algorithm divides the text into chunks, each containing 1024 words. The size of these chunks is chosen to balance between maintaining enough context for meaningful summaries and keeping the chunks manageable for processing by summarization models.

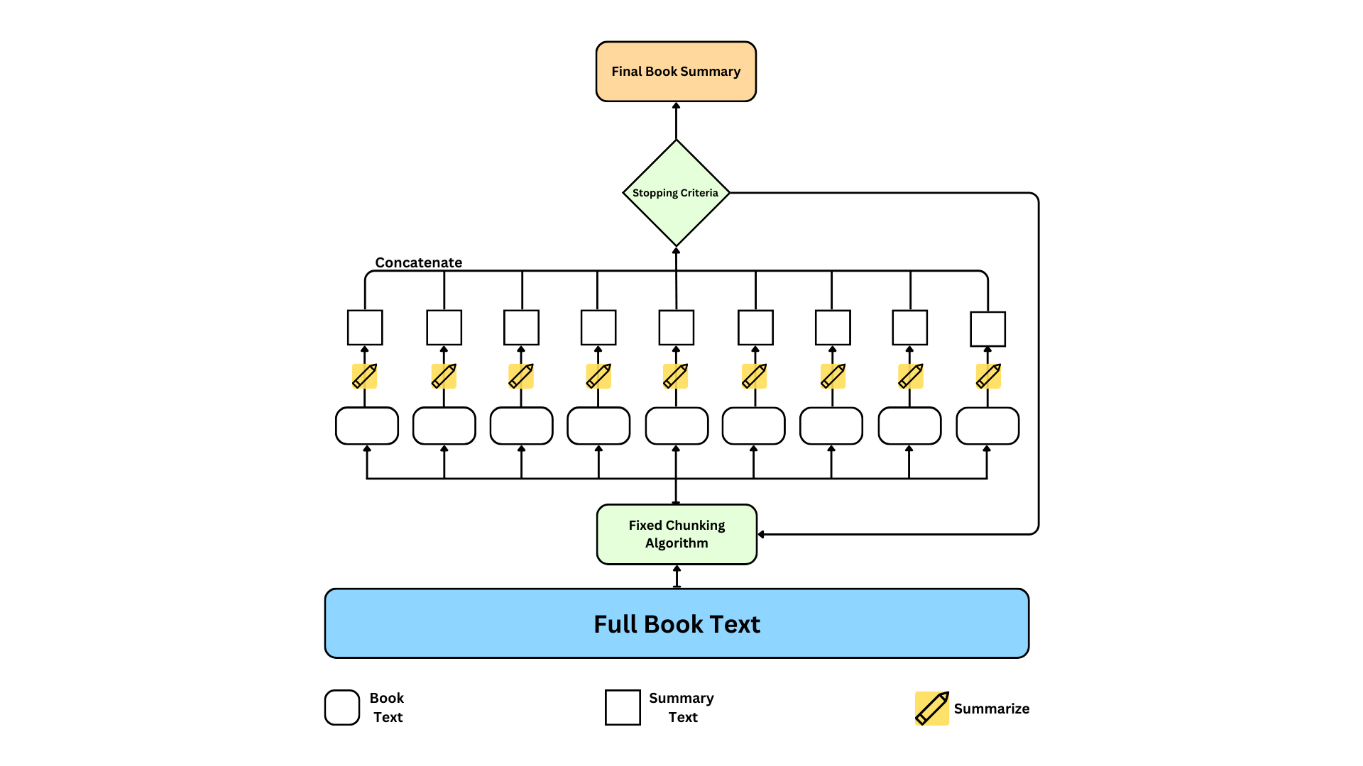
**Chunks Summarization:**

Each 1024-word chunk is then individually summarized using an abstractive summarization model. This model is trained to generate coherent and concise summaries that capture the essential information of each chunk while preserving the narrative flow and key details.

The resulting summaries of these chunks are then concatenated to form a new, intermediate text. This intermediate text undergoes the chunking process again, where it is divided into new chunks, each of 1024 words. These chunks are again summarized, and the process of concatenation and re-chunking is repeated recursively.

This recursive summarization continues until a stopping criterion is met. Specifically, the process stops when the total number of chunks, each now containing summaries, is equal to or less than five. This stopping criterion ensures that the final set of summaries is concise enough to be easily comprehensible while still capturing the critical elements of the original narrative.

By employing this recursive task decomposition approach, we can efficiently and effectively summarize long narrative novels, producing summaries that are both manageable in length and rich in content. Our method not only simplifies the summarization process but also enhances the quality of the final summaries, making them suitable for various applications, such as quick reviews, academic analysis, and content recommendation



**Fig 1: Overview of the Recursive Task Decomposition Pipeline**

Figure 1 illustrates the detailed process of our recursive task decomposition pipeline for summarizing long narrative novels. This pipeline is designed to handle extensive and complex text by breaking it down into manageable chunks and iteratively summarizing these chunks to produce a concise final summary.

## **3.3** **Dataset**

We have used the BookSum Dataset ‎[10] to fine-tune our transformer model. *BookSum is a collection of datasets for long-form narrative summarization that focuses on literature, including novels, plays, and stories. It is designed to challenge summarization systems with complex, long-range causal and temporal dependencies, rich discourse structures, and the need for highly abstractive summaries. The dataset consists of documents provided with human-written summaries at three levels of granularity: paragraph-level, chapter-level, and book-level.*

#### **Key Features:**

1. **Source Domain**: The dataset includes literary works from the Project Gutenberg repository, focusing on the public domain.
2. **Summary Types**: Highly abstractive human-written summaries are provided at paragraph, chapter, and book levels, each with increasing difficulty.
3. **Challenges**: The dataset introduces challenges such as processing long documents, understanding complex dependencies, handling rich discourse structures, and generating abstractive summaries.
4. **Size and Scope**: The dataset contains 146,532 paragraph-level examples, 12,630 chapter-level examples, and 405 book-level examples, making it one of the largest and most detailed datasets for this purpose.

**Data Collection and Processing:**

Data Sources: Documents were collected from Project Gutenberg, while summaries were sourced from the Web Archive.

**Data Cleaning and Splitting**: Texts were cleaned to remove metadata and were split into individual chapters and paragraphs. Automatic and manual methods ensured high-quality data alignment between texts and summaries.

**Alignment**: Coarse-grained and fine-grained alignments were performed to pair texts with summaries, involving both automated tools and manual inspection.

In our project we utilized the BookSum dataset, specifically focusing on paragraph-level summarization to fine-tune our model. This dataset is well-suited for our needs as it provides highly abstractive human-written summaries of literary works from the Project Gutenberg repository, emphasizing long-range dependencies and rich discourse structures. The challenges presented by BookSum, including the complexity of summarizing extensive texts and the requirement for abstractive summaries, align with our project's goals of generating concise yet comprehensive summaries. Additionally, the availability of summaries at different levels of granularity (paragraph, chapter, book) allows for flexible experimentation. By focusing on the paragraph level, we effectively break down complex narratives into manageable sections. This approach, coupled with the dataset's support for benchmarking and baseline comparisons, makes BookSum an ideal choice for advancing the state of the art in book summarization.

**3.4 Model:**

In the context of summarizing long narrative novels, the need for a model that can process text quickly and accurately is paramount. The chosen model for this task is the T5 (Text-to-Text Transfer Transformer) ‎[11], which has demonstrated impressive capabilities in various natural language processing (NLP) tasks.

**Model Overview**

The T5 model is designed to handle a wide range of text-based tasks by converting all of them into a text-to-text format. This approach allows for a unified framework where the model takes text as input and generates text as output, making it particularly suitable for tasks like summarization, translation, and question answering.

**Architecture**

The T5 model is based on the Transformer architecture, introduced by Vaswani et al. (2017) ‎[12]. The Transformer architecture consists of an encoder-decoder structure, where both the encoder and decoder are composed of multiple layers of self-attention mechanisms and feed-forward neural networks.

***Encoder***: The encoder processes the input text by mapping it to a continuous representation. It consists of a stack of identical layers, each containing a self-attention mechanism and a feed-forward neural network.

***Decoder***: The decoder generates the output text by attending to the encoder's output and the previously generated tokens. Like the encoder, it is composed of multiple layers with self-attention and feed-forward networks.

**Pre-training**

The model is pre-trained on a large, diverse dataset using unsupervised learning. This stage helps the model develop general-purpose knowledge about language. For the T5 model, the pre-training objective is to reconstruct corrupted input text, where parts of the text are masked, and the model learns to predict the missing tokens.

**Why T5?**

The choice of the T5 model for this research is based on several factors:

**Flexibility:** The text-to-text framework of T5 allows for seamless adaptation to various NLP tasks, including summarization.

**Performance:** T5 has achieved state-of-the-art results in multiple benchmarks, demonstrating its effectiveness in generating high-quality text.

Scalability: The model's architecture supports training on large datasets and can handle long input sequences, making it ideal for summarizing lengthy novels.

The T5 model's robust architecture, comprehensive training approach, and demonstrated performance make it a suitable choice for the task of summarizing long narrative novels. Its ability to transform text into meaningful and concise summaries aligns well with the goals of this research.

**3.5 Fine-Tuning:**

The fine-tuning of the T5 model was conducted using the BookSum Dataset, specifically focusing on paragraph-level summaries.

***Data Split:***

The dataset was split into two parts: 80% for training and 10% for validation and 10% for Testing. This split ensures that the model has a substantial amount of data to learn from while also providing a separate set for evaluating its performance.

**Fine-Tuning Procedure**

The fine-tuning process involves adjusting the pre-trained T5 model's weights to optimize its performance on the specific task of summarizing paragraphs. The model is trained to minimize the cross-entropy loss between the predicted and actual summaries. Cross-entropy loss measures the difference between the predicted probability distribution and the actual distribution, guiding the model to make accurate predictions.

The loss calculation and weight adjustment during the fine-tuning process of the T5 model involve the following steps:

1. **Loss Calculation:**

***Objective:*** The objective during fine-tuning is to minimize the loss function, which in the case of T5, is typically the cross-entropy loss. This loss measures the difference between the predicted token probabilities and the actual tokens in the target sequence.

***Process:*** During fine-tuning, the model's predictions are compared to the ground truth labels (i.e., the actual text it should produce). The cross-entropy loss function calculates the error by considering the predicted probabilities of the correct tokens.

1. **Weight Adjustment:**

***Backpropagation:*** Once the loss is calculated, the backpropagation algorithm is used to compute the gradients of the loss with respect to the model parameters (weights and biases). These gradients indicate how much the loss would change with a small change in the parameters.

***Gradient Descent:*** The model parameters are then updated using gradient descent or one of its variants (e.g., Adam optimizer). In each iteration, the parameters are adjusted in the direction that reduces the loss. Specifically, the weights are updated as follows:

## where represents the model parameters, is the learning rate, and is the gradient of the loss with respect to the parameters.

By iteratively performing the fine-tuning steps, the T5 model learns to produce accurate and coherent summaries for paragraphs in the BookSum Dataset. This fine-tuning process allows the model to adapt its general language understanding capabilities to the specific nuances of long narrative summarization.

**IV. Results and Experiments**

**4.1 Rouge:**

ROUGE ‎[6], or Recall-Oriented Understudy for Gisting Evaluation, compares a model-generated summary to a "gold-standard" human-generated summary. It is the widely applied metric for evaluating the performance of NLP models. Given that summary generation is subjective, ROUGE measures how well the model-generated summary matches the human-generated one.

ROUGE-N evaluates summaries using an N-gram recall metric, comparing the number of N-grams (contiguous sequences of n items) in the model summary to those in the human summary. Variants of the ROUGE metric include ROUGE-1 (unigram), ROUGE-2 (bigram), ROUGE-3 (trigram), and ROUGE-L (Longest Common Subsequence).

ROUGE calculates precision, recall, and F-score based on these comparisons. The proposed model uses ROUGE-N and obtained the following results for the given application.

**4.2 T5 Evaluation:**

| Score | Rouge | | |
| --- | --- | --- | --- |
| Rouge-L | Rouge-2 | Rouge-1 |
| F1-Score | 0.175 | 0.05 | 0.196 |
| Precision | 0.158 | 0.045 | 0.176 |
| Recall | 0.229 | 0.069 | 0.06 |

**Table 1**

The T5 model was fine-tuned for 35 epochs, and its performance was evaluated using test data distinct from the data used for training and validation. The results, summarized in Table 1, demonstrate the model's performance across different ROUGE metrics.

These results indicate the effectiveness of the T5 model in generating summaries, with varying degrees of precision, recall, and F1-score across different ROUGE metrics. The evaluation based on separate test data (10% of the data for testing) ensures that the model's performance reflects its ability to generalize beyond the training and validation sets.

**4.3 Experiments:**

This section outlines the experiments conducted to compare the performance of the BART ‎[13] and T5 models in the domain of book summarization. The experiments aimed to evaluate the models both in their zero-shot capabilities and after fine-tuning on a specialized dataset. The evaluation employed the ROUGE metric, a standard for assessing the quality of text summarization.

The experiments were structured into two phases: zero-shot evaluation and fine-tuning.

**Zero-Shot Evaluation:**

**Models:** BART and T5

**Objective:** To assess the performance of the pre-trained models without any task-specific fine-tuning.

**Dataset:** A set of 19 novels with human-written summaries.

**Evaluation Metric:** ROUGE-1, ROUGE-2, and ROUGE-L.

|  |  |  |
| --- | --- | --- |
|  | BART | T5 |
| Rouge-L (F1-Score) | 0.155 | 0.113 |

**Table 2**

**Analysis:** BART outperformed T5 in the zero-shot evaluation, indicating that BART's pre-trained weights are more effective for summarization tasks without additional

**Fine-Tuning:**

Models: BART and T5

**Objective:** To fine-tune the models on the BookSum dataset and evaluate their performance improvements.

**Dataset:** The same set of 19 novels with human-written summaries.

**Fine-Tuning Process:** Both models were fine-tuned for several epochs using Kaggle's GPU infrastructure.

**Evaluation Metric:** ROUGE-1, ROUGE-2, and ROUGE-L.

|  |  |  |
| --- | --- | --- |
|  | BART | T5 |
| Rouge-L (F1-Score) | 0.17 | 0.17 |

**Table 3**

**Analysis:** After fine-tuning, T5 showed significant improvements and matched the performance of BART. Despite both models achieving similar accuracy post-fine-tuning, T5's smaller size led to faster inference times, making it a more efficient choice.

**Discussion:**

BART: BART's denoising autoencoder setup proves beneficial for generating coherent summaries in a zero-shot context. However, its performance gain from fine-tuning is less pronounced compared to T5.

T5: T5's text-to-text transfer transformer framework allows it to effectively learn from fine-tuning data, resulting in superior performance post-fine-tuning. T5's ability to generalize from specific fine-tuning tasks significantly enhances its summarization quality. Additionally, T5's smaller model size results in faster processing times, further supporting its use over BART for efficient summarization tasks.

**Experiments Conclusion:**

The experiments underscore the relative strengths and weaknesses of the BART and T5 models for book summarization:

BART: Exhibits superior performance in zero-shot summarization but demonstrates moderate improvements after fine-tuning.

T5: Initially performs lower in zero-shot settings but surpasses BART significantly post-fine-tuning, with the added advantage of faster processing times due to its smaller size.

These findings suggest that for tasks requiring high-quality summarization with available task-specific data, fine-tuning T5 is the preferable approach. Conversely, for applications where fine-tuning is not feasible, BART's zero-shot capabilities make it a strong candidate. This insight is crucial for selecting the appropriate model based on the specific requirements and constraints of book summarization projects.

# **V. Conclsions**

In this paper, we have explored the application of the T5 model in the domain of novel long narrative book summarization. Beginning with an overview of the Transformer architecture and the specific adaptation of T5, we delved into its fine-tuning process using the BookSum Dataset at the paragraph level. Our findings illustrate the effectiveness of T5 in capturing and summarizing the essence of lengthy narratives, facilitated by its ability to adapt to domain-specific data through extensive training epochs.

Through evaluation on distinct test datasets, we validated the model's robustness and generalizability, highlighting its capability to generate coherent and informative summaries that preserve the original narrative's key elements. The results not only demonstrate the viability of leveraging advanced NLP models like T5 for complex summarization tasks but also underscore the importance of dataset curation and fine-tuning strategies in optimizing performance.

Looking forward, the integration of such technologies holds promise for enhancing accessibility to literary content, aiding in educational settings, and potentially transforming how narratives are consumed and analyzed in digital formats. As advancements continue, further exploration into multi-modal approaches and broader corpora could yield even more nuanced and contextually rich summaries, pushing the boundaries of narrative understanding and computational linguistics.

In conclusion, while challenges remain in balancing fidelity with brevity in summarization tasks, the T5 model represents a significant step towards automated narrative compression, offering new avenues for exploration in the evolving landscape of natural language processing and literary analysis.

**References**

1. Ladhak, F., Li, B., Al-Onaizan, Y., & McKeown, K. (2020, May 4). Exploring content selection in summarization of novel chapters. arXiv.org. <https://arxiv.org/abs/2005.01840>
2. Radev, D., Allison, T., Blair-Goldensohn, S., Blitzer, J., Çelebi, A., Dimitrov, S., Drabek, E. F., Hakim, A., Lam, W., Liu, D., Otterbacher, J., Qi, H., Saggion, H., Teufel, S., Topper, M., Winkel, A., and Zhang, Z. (2004). Mead-a platform for multidocument multilingual text summarization. In Proceedings of the Fourth International Conference on Language Resources and Evaluation.
3. Kazantseva, A. (2006). An approach to summarizing short stories. In Student Research Workshop.
4. Porwal, S., Bewoor, L., & Deshpande, V. (2022). Transformer Based Implementation for Automatic Book Summarization1International Journal of Intelligent Systems and Applications in Engineering, 10(3s), 123–12823.
5. Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2018, October 11). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv.org. <https://arxiv.org/abs/1810.04805>
6. Lin, C.-Y. and Och, F. J. (2004). Automatic evaluation of machine translation quality using longest common subsequence and skip-bigram statistics. In Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics, page 605. Association for Computational Linguistics.
7. Wu, J., Ouyang, L., Ziegler, D. M., Stiennon, N., Lowe, R., Leike, J., & Christiano, P. (2021, September 22). Recursively Summarizing Books with Human Feedback. arXiv.org. <https://arxiv.org/abs/2109.10862>
8. Christiano, P., Leike, J., Brown, T. B., Martic, M., Legg, S., & Amodei, D. (2017, June 12). Deep reinforcement learning from human preferences. arXiv.org. https://arxiv.org/abs/1706.03741
9. Leike, J., Krueger, D., Everitt, T., Martic, M., Maini, V., and Legg, S. (2018). Scalable agent alignment via reward modeling: a research direction. arXiv preprint arXiv:1811.07871.
10. Krysci ´ nski, W., Rajani, N., Agarwal, D., Xiong, C., and Radev, D. (2021). Booksum: A collection of ´ datasets for long-form narrative summarization. arXiv preprint arXiv:2105.08209.
11. Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., and Liu, P. J. (2019). Exploring the limits of transfer learning with a unified text-to-text transformer. arXiv preprint arXiv:1910.10683.
12. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017, June 12). Attention is all you need. arXiv.org. <https://arxiv.org/abs/1706.03762>
13. Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2019, October 29). BART: Denoising Sequence-to-Sequence Pre-training for natural language generation, Translation, and Comprehension. arXiv.org. https://arxiv.org/abs/1910.13461

**Examples**

|  |
| --- |
| **Paragraph 1 (Animal Farm Book)** |
| With the ring of light from his lantern dancing from side to side, he lurched across the yard, kicked off his boots at the back door, drew himself a last glass of beer from the barrel in the scullery, and made his way up to bed, where Mrs. Jones was already snoring. As soon as the light in the bedroom went out there was a stirring and a fluttering all through the farm buildings. Word had gone round during the day that old Major, the prize Middle White boar, had had a strange dream on the previous night and wished to communicate it to the other animals. It had been agreed that they should all meet in the big barn as soon as Mr. Jones was safely out of the way. Old Major (so he was always called, though the name under which he had been exhibited was Willingdon Beauty) was so highly regarded on the farm that everyone was quite ready to lose an hour’s sleep in order to hear what he had to say. At one end of the big barn, on a sort of raised platform, Major was already ensconced on his bed of straw, under a lantern which hung from a beam. He was twelve years old and had lately grown rather stout, but he was still a majestic-looking pig, with a wise and benevolent appearance in spite of the fact that his tushes had never been cut. Before long the other animals began to arrive and make themselves comfortable after their different fashions. First came the three dogs, Bluebell, Jessie, and Pincher, and then the pigs, who settled down in the straw immediately in front of the platform. The hens perched themselves on the window-sills, the pigeons fluttered up to the rafters, the sheep and cows lay down behind the pigs and began to chew the cud. The two cart-horses, Boxer and Clover, came in together, walking very slowly and setting down their vast hairy hoofs with great care lest there should be some small animal concealed in the straw. Clover was a stout motherly mare approaching middle life, who had never quite got her figure back after her fourth foal. Boxer was an enormous beast, nearly eighteen hands high, and as strong as any two ordinary horses put together. A white stripe down his nose gave him a somewhat stupid appearance, and in fact he was not of first-rate intelligence, but he was universally respected for his steadiness of character and tremendous powers of work. After the horses came Muriel, the white goat, and Benjamin, the donkey. Benjamin was the oldest animal on the farm, and the worst tempered. He seldom talked, and when he did, it was usually to make some cynical remark – for instance, he would say that God had given him a tail to keep the flies off, but that he would sooner have had no tail and no flies. Alone among the animals on the farm he never laughed. If asked why, he would say that he saw nothing to laugh at.Nevertheless, without openly admitting it, he was devoted to Boxer; the two of them usually spent their Sundays together in the small paddock beyond the orchard, grazing side by side and never speaking. The two horses had just lain down when a brood of ducklings, which had lost their mother, filed into the barn, cheeping feebly and wandering from side to side to find some place where they would not be trodden on. Clover made a sort of wall round them with her great foreleg, and the ducklings nestled down inside it and promptly fell asleep. At the last moment Mollie, the foolish, pretty white mare who drew Mr. Jones’s trap, came mincing daintily in, chewing at a lump of sugar. She took a place near the front and began flirting her white mane, hoping to draw attention to the red ribbons it was plaited with. Last of all came the cat, who looked round, as usual, for the warmest place, and finally squeezed herself in between Boxer and Clover; there she purred contentedly throughout Major’s speech without listening to a word of what he was saying. All the animals were now present except Moses, the tame raven, who slept on a perch behind the back door. When Major saw that they had all made themselves comfortable and were waiting attentively, he cleared his throat and began: ‘Comrades, you have heard already about the strange dream that I had last night. But I will come to the dream later. I have something else to say first. |
| **Summary 1** |
| Major, the prize Middle White boar, has had a strange dream the previous night, and he wants to tell it to the other animals., The other animals arrive in the big barn, and Major is ensconced on his bed of straw, under a lantern hung from a beam., Boxer and Clover come in together, walking very slowly and setting down their vast hairy hoofs with great care lest there should be some small animal concealed in the straw., Muriel, the white goat, and Benjamin, the donkey., Mollie, the foolish, pretty white mare who drew Mr. Joness trap, came mincing daintily throughout Majors speech without listening a word of what he was talking about., Moses, the tame raven, slept on a perch behind the back door., All the animals arrived and made themselves comfortable after their different fashions. |

|  |
| --- |
| **Paragraph 2 (Animal Farm Book)** |
| I do not think, comrades, that I shall be with you for many months longer, and before I die, I feel it my duty to pass on to you such wisdom as I have acquired. I have had a long life, I have had much time for thought as I lay alone in my stall, and I think I may say that I understand the nature of life on this earth as well as any animal now living. It is about this that I wish to speak to you. ‘Now, comrades, what is the nature of this life of ours? Let us face it: our lives are miserable, laborious, and short. We are born, we are given just so much food as will keep the breath in our bodies, and those of us who are capable of it are forced to work to the last atom of our strength; and the very instant that our usefulness has come to an end we are slaughtered with hideous cruelty. No animal in England knows the meaning of happiness or leisure after he is a year old. No animal in England is free. The life of an animal is misery and slavery: that is the plain truth. ‘But is this simply part of the order of nature? Is it because this land of ours is so poor that it cannot afford a decent life to those who dwell upon it? No, comrades, a thousand times no! The soil of England is fertile, its climate is good, it is capable of affording food in abundance to an enormously greater number of animals than now inhabit it. This single farm of ours would support a dozen horses, twenty cows, hundreds of sheep – and all of them living in a comfort and a dignity that are now almost beyond our imagining. Why then do we continue in this miserable condition? Because nearly the whole of the produce of our labour is stolen from us by human beings. There, comrades, is the answer to all our problems. It is summed up in a single word – Man. Man is the only real enemy we have. Remove Man from the scene, and the root cause of hunger and overwork is abolished for ever. ‘Man is the only creature that consumes without producing. He does not give milk, he does not lay eggs, he is too weak to pull the plough, he cannot run fast enough to catch rabbits. Yet he is lord of all the animals. He sets them to work, he gives back to them the bare minimum that will prevent them from starving, and the rest he keeps for himself. Our labour tills the soil, our dung fertilises it, and yet there is not one of us that owns more than his bare skin. You cows that I see before me, how many thousands of gallons of milk have you given during this last year? And what has happened to that milk which should have been breeding up sturdy calves? Every drop of it has gone down the throats of our enemies. And you hens, how many eggs have you laid in this last year, and how many of those eggs ever hatched into chickens? The rest have all gone to market to bring in money for Jones and his men. And you, Clover, where are those four foals you bore, who should have been the support and pleasure of your old age? Each was sold at a year old – you will never see one of them again. In return for your four confinements and all your labour in the fields, what have you ever had except your bare rations and a stall? ‘And even the miserable lives we lead are not allowed to reach their natural span. For myself I do not grumble, for I am one of the lucky ones. I am twelve years old and have had over four hundred children. Such is the natural life of a pig. But no animal escapes the cruel knife in the end. You young porkers who are sitting in front of me, every one of you will scream your lives out at the block within a year. To that horror we all must come – cows, pigs, hens, sheep, everyone. Even the horses and the dogs have no better fate. You, Boxer, the very day that those great muscles of yours lose their power, Jones will sell you to the knacker, who will cut your throat and boil you down for the foxhounds. As for the dogs, when they grow old and toothless, Jones ties a brick round their necks and drowns them in the nearest pond. |
| **Summary 2** |
| Man is the only real enemy we have, he says, and the root cause of hunger and overwork is abolished forever., He says that the land of England is fertile, it is capable of food in abundance to an enormously greater number of animals than now inhabit it, and that the soil of England has fertile soil, but there is not one of us that owns more than his bare skin., No animal in England knows the meaning of happiness or leisure after he is a year old., The life of an animal is misery and slavery: that is the plain truth., There is the answer to all our problems. It is summed up in a single word - Man., We are born, we are given just so much food as will keep the breath in our bodies, and those of us who are capable of it are forced to work to the lastatom of our strength; and the very instant that our usefulness has come to an end we are slaughtered with hideous cruelty., Every drop of that milk has gone down the throats of our enemies., And you, Boxer, I do not grumble, for I am one of the lucky ones. |

|  |
| --- |
| **Paragraph 3** **(Animal Farm Book)** |
| ‘Is it not crystal clear, then, comrades, that all the evils of this life of ours spring from the tyranny of human beings? Only get rid of Man, and the produce of our labour would be our own. Almost overnight we could become rich and free. What then must we do? Why, work night and day, body and soul, for the overthrow of the human race! That is my message to you, comrades: Rebellion! I do not know when that Rebellion will come, it might be in a week or in a hundred years, but I know, as surely as I see this straw beneath my feet, that sooner or later justice will be done. Fix your eyes on that, comrades, throughout the short remainder of your lives! And above all, pass on this message of mine to those who come after you, so that future generations shall carry on the struggle until it is victorious. ‘And remember, comrades, your resolution must never falter. No argument must lead you astray. Never listen when they tell you that Man and the animals have a common interest, that the prosperity of the one is the prosperity of the others. It is all lies. Man serves the interests of no creature except himself. And among us animals let there be perfect unity, perfect comradeship in the struggle. All men are enemies. All animals are comrades.’ At this moment there was a tremendous uproar. While Major was speaking four large rats had crept out of their holes and were sitting on their hindquarters, listening to him. The dogs had suddenly caught sight of them, and it was only by a swift dash for their holes that the rats saved their lives. Major raised his trotter for silence. ‘Comrades,’ he said, ‘here is a point that must be settled. The wild creatures, such as rats and rabbits – are they our friends or our enemies? Let us put it to the vote. I propose this question to the meeting: Are rats comrades?’ The vote was taken at once, and it was agreed by an overwhelming majority that rats were comrades. There were only four dissentients, the three dogs and the cat, who was afterwards discovered to have voted on both sides. Major continued: ‘I have little more to say. I merely repeat, remember always your duty of enmity towards Man and all his ways. Whatever goes upon two legs is an enemy. Whatever goes upon four legs, or has wings, is a friend. And remember also that in fighting against Man, we must not come to resemble him. Even when you have conquered him, do not adopt his vices. No animal must ever live in a house, or sleep in a bed, or wear clothes, or drink alcohol, or smoke tobacco, or touch money, or engage in trade. All the habits of Man are evil. And, above all, no animal must ever tyrannise over his own kind. Weak or strong, clever or simple, we are all brothers. No animal must ever kill any other animal. All animals are equal. ‘And now, comrades, I will tell you about my dream of last night. I cannot describe that dream to you. It was a dream of the earth as it will be when Man has vanished. But it reminded me of something that I had long forgotten. Many years ago, when I was a little pig, my mother and the other sows used to sing an old song of which they knew only the tune and the first three words. I had known that tune in my infancy, but it had long since passed out of my mind. Last night, however, it came back to me in my dream. And what is more, the words of the song also came back – words, I am certain, which were sung by the animals of long ago and have been lost to memory for generations. I will sing you that song now, comrades. I am old and my voice is hoarse, but when I have taught you the tune, you can sing it better for yourselves. It is called Beasts of England.’ Old Major cleared his throat and began to sing. As he had said, his voice was hoarse, but he sang well enough, and it was a stirring tune, something between Clementine and La Cucaracha. The words ran: Beasts of England, beasts of Ireland, Beasts of every land and clime, Hearken to my joyful tidings Of the golden future time. Soon or late the day is coming, Tyrant Man shall be o’erthrown, And the fruitful fields of England Shall be trod by beasts alone. Rings shall vanish from our noses, And the harness from our back, Bit and spur shall rust forever, Cruel whips no more shall crack. |
| **Summary 3** |
| In a dream, he recalls a song he heard as a young pig, which reminds him of an old song that he learned when he was a little pig., He sings the song Beasts of England, which he has memorized for years. |