

TEXT SUMMARIZATION

NLP project

By:

Pratyush Kumar Singh[RA2211027010088]

K Yogesh[RA2211027010099]

R Kaviya[RA2211027010101]



INTRODUCTION

What is Text Summarization?

- Text summarization is the process of automatically condensing a long document into a shorter version while retaining key information.
- It helps extract essential insights from large volumes of data efficiently.

Why is it Important?

- In today's fast-paced digital world, people don't have time to read lengthy articles and reports.
- Automated summarization helps improve comprehension, efficiency, and productivity.
- Used in news summarization, research paper analysis, legal documentation, and content recommendation.



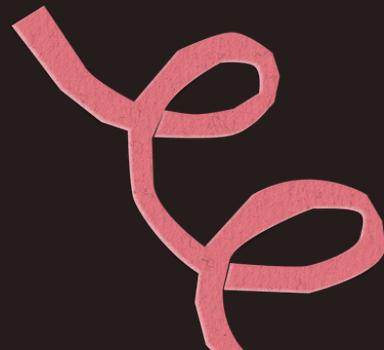
ABSTRACT

- This project explores the use of Natural Language Processing (NLP) techniques for automated text summarization.
- Extractive summarization selects key sentences from the text using techniques like Latent Semantic Analysis (LSA).
- Abstractive summarization is implemented using deep learning-based transformer models (BART, T5).
- The aim is to enhance information retrieval and improve readability by generating high-quality summaries.
- The project evaluates different summarization approaches using ROUGE scores and human evaluation metrics.

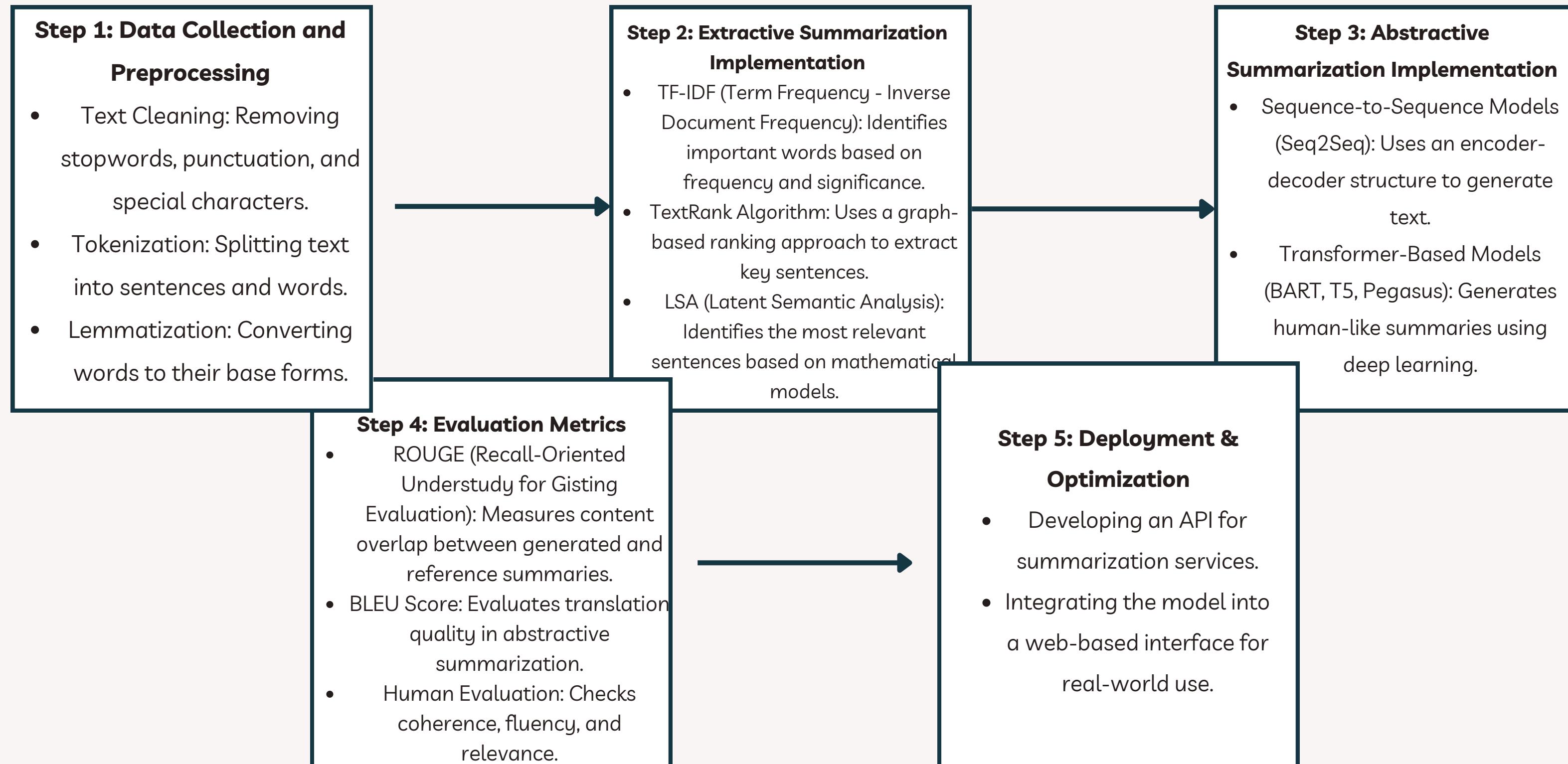


PROBLEM STATEMENT

- **Challenges in Manual Summarization:**
 - Time-consuming and labor-intensive.
 - Subjective bias in summary generation.
 - Difficulty in identifying key information in large datasets.
- **Limitations of Existing Automated Summarization Methods:**
 - Extractive methods often produce disjointed summaries.
 - Abstractive models sometimes generate inaccurate or hallucinated content.
- **Need for NLP-driven Summarization:**
 - Improves efficiency and accuracy in text summarization.
 - Can be used in various industries such as journalism, healthcare, finance, and legal sectors.



METHODOLOGY



SOLUTION & IMPLEMENTATION

- **Extractive Summarization Approach:**

1. Implemented using the Sumy library and Latent Semantic Analysis (LSA).
2. Extracts key sentences from the document.
3. Generates summaries quickly but lacks coherence.

- **Abstractive Summarization Approach:**

4. Implemented using Hugging Face's Transformers (BART, T5).
5. Generates human-like summaries with improved readability.
6. More computationally expensive than extractive methods.

- **Comparison of Results:**

7. Extractive Summarization: Retains original sentences but may sound robotic.
8. Abstractive Summarization: Generates a more fluid and natural summary.

Applications of Text Summarization

- News Aggregation (Google News, Inshorts)
- Search Engines (Google Snippets, Bing)
- Chatbots & Customer Support
- Legal & Financial Document Summarization
- Scientific & Medical Research Papers

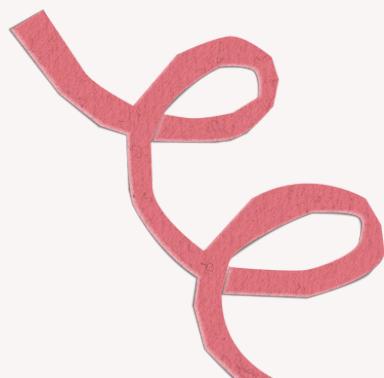
CONCLUSION & FUTURE WORK

Conclusion:

- NLP-based text summarization significantly improves efficiency and content readability.
- Extractive summarization is faster and computationally cheaper, but lacks flexibility.
- Abstractive summarization provides high-quality summaries, but requires more training data and processing power.

Future Enhancements:

- Improving factual accuracy in abstractive summarization.
- Exploring hybrid approaches combining extractive & abstractive methods.
- Multilingual summarization for global applications.
- Real-time summarization for streaming data (e.g., live news, social media).
- Integration with voice assistants (Alexa, Siri) and mobile apps.



THANK YOU!