Assam Engineering College
Department of Electronics and Telecommunication Engineering

FINAL YEAR PROJECT PRESENTATION CUM VIVA

SUMMARY GENERATOR WEB APPLICATION USING AI

PRESENTED BY: (GROUP NO. 5)

Kuladeep Barman (190610026024)
Srinjana Deb (190610026042)
Salik Ahmed (190612826004)
Princi Priya Gogoi (200650026002)

UNDER THE GUIDANCE OF:

Siddhanta Borah (Asst. Professor, ETE, AEC)
Dr. Mukut Senapati (Asst. Professor, ETE, AEC)

CONTENTS

PROBLEM STATEMENT

The tremendous abundance of material available on the internet has produced an odd paradox:

People are immersed in information, yet they are yearning for wisdom. It is tough to keep up with the internet's daily production of billions of articles on any given topic. In order to absorb information more effectively in this case, one needs to devote more reading time. But this is not helpful in today's scenario where people are constantly hustling.

So, what can be done?

A SOLUTION

Summary Generator Web Application is a user-friendly and convenient tool designed to automatically generate summaries from input text or documents. It utilizes advanced natural language processing techniques and machine learning algorithms to extract key information and produce concise summaries that capture the essence of the original content.

INTRODUCTION

Text summarization is the process of making a synopsis from a given text document while keeping the important information and its meaning. Automatic summarization has become an essential method for accurately locating significant information in vast amounts of text in a short amount of time and with minimal effort.

Text summarization is one of the Natural Language Processing (NLP) applications that will undoubtedly have a significant influence on our lives. Automatic Text Summarization is among the most complex and intriguing topics in (NLP). It is the way of forming a brief and coherent summary of writing from a variety of text sources, including books, news stories, blog posts, research papers, emails, and tweets. The advent of vast volumes of textual data is driving up demand for automatic summarization technologies.

APPROACH TO TEXT SUMMARISATION

EXTRACTIVE SUMMARISATION

STEPS OF EXTRACTIVE SUMMARISATION:

Let us take a paragraph where we perform the following steps :

STEP 1	First, split the paragraph into n correspondence sentences.
STEP 2	Text Processing
STEP 3	Tokenization
STEP 4	Evaluate the weighted occurrence of frequency of the word
STEP 5	Substitute the word with their weighted frequency

THE EXAMPLE

The paragraph is as follows:

Peter and Elizabeth took a taxi to attend the night party in the city. while in the party, Elizabeth collapsed and was rushed to the hospital. Since she was diagnosed with a brain injury, the doctor told Peter to stay besides her until she gets well. Peter stayed with her at the hospital for 3 days without leaving.

IMPLEMENTING EXTRACTION-BASED SUMMARISATION

STEP 1 Break the paragraph into sentences

1: "Peter and Elizabeth took a taxi to attend the night party in the city."
2: "While in the party, Elizabeth collapsed and was rushed to the hospital."
3: "Since she was diagnosed with a brain injury, the doctor told Peter to stay besides her until she
gets well. "
4: "Peter stayed with her at the hospital for 3 days without leaving"

STEP 2 Text processing

Text processing is done by removing extremely common words.

Result of the text processing:
1: Peter Elizabeth took taxi attend night party city.
2: party Elizabeth collapsed rush hospital.
3. diagnose brain injury doctor told Peter stay beside get well
4. Peter stayed hospital without leaving.

STEP 3 Tokenization

Tokenization is done to get all the words present in the sentences.

["Peter", "Elizabeth", "took"....., "leaving"]

STEP 4

Calculation for weighted occurence

1. Peter and Elizabeth took a taxi to attend the night party in the city.	1+0.67+0.33+0.33+0.33+0.67+0.33=3.99
2. while in the party, Elizabeth collapsed and was rushed to the hospital.	0.67+0.67+0.33+0.33+067=2.67
3. since she was diagonised with a brain injury ,the doctor tolds peter to stay besides her while she gets well.	3.97
4. therefore peter stayed with her for three days without leaving	3.33

So the first sentence carries the most weight in the paragraph. Therefore it can give the best representation of the paragraph. If the first sentence is combined with the third sentence in the paragraph then a better summary can be generated.

This is the basic illustration of how to generate an extraction-based text summarization.

ABSTRACTIVE SUMMARISATION

STEPS OF ABSTRACTIVE SUMMARISATION:

PRE-PROCESSING

UNDERSTANDING THE TEXT

ENCODING THE TEXT

MODEL TRAINING

SUMMARISATION GENERATION

POST-PROCESSING

EVALUATION AND REFINEMENT

BLEU SCORE AND ROUGE

Bleu measures precision: how much the words (and/or n-grams) in the machine-generated summaries appeared in the human reference summaries.

Precision = Number of important sentences/Total number of sentences summarized.

Rouge measures recall: how much the words (and/or n-grams) in the human reference summaries appeared in the machine-generated summaries.

Recall = Total number of Important sentences Retrieved / Total number of important sentences present.

F1 MEASURE

The F1-measure, also called the F1-score, is a measure of a model's accuracy on a dataset. The F-score is a way of combining the precision and recall of the model, and it is defined as the harmonic mean of the model's precision and recall.

F Score = 2 x (Bleu x Rouge) / (Bleu + Rouge)
= Total number of words in the summary / Total number of words in original document.

EVALUATION OF OUR MODEL

CONCLUSION

However, abstractive summarization is a challenging task as it requires language understanding, coherence, and generating grammatically correct and meaningful sentences. The quality of abstractive summaries can vary depending on the complexity of the text and the performance of the underlying algorithms.

Ongoing research and advancements in natural language processing are continuously improving the accuracy and effectiveness of abstractive summarization techniques. Abstractive summarization has the potential to enhance information retrieval, aid in decision-making, and improve the accessibility of large volumes of textual data.

THANK YOU!