

NEWS PAPER SUMMARIZATION ALONG WITH NEWS VIDEOS **ON YOUTUBE**

Abstract:

As online news consumption continues shifting towards video platforms, viewers face rising challenges in distilling key details from lengthy transcripts. This project aims to enhance the news digestion experience by utilizing recent advances in natural language processing to automatically generate concise summaries of YouTube news videos.

At the core, two neural sequence-to-sequence models are employed - BART and T5 - for their proficiency on text summarization tasks. Both incorporate encoder-decoder architectures, interpreting input sequences and predicting condensed summaries reflecting salient information. However, BART uniquely combines bidirectional context modeling enabling richer understanding compared to conventional left-to-right approaches. Meanwhile, T5 maintains a consistent text-in to text-out framework making adapting between datasets and tasks more seamless. The models are trained on an amalgamated dataset blending YouTube transcripts and the CNN/DailyMail corpus. This multi-source input aids model resilience to linguistic variances in topics and styles. Summarization quality is evaluated via the widely adopted ROUGE metric which quantifies overlap with ground truth summaries across factors like word choice and order. For ease-of-use, a Flask web application allows users to simply submit a YouTube link and view a machine-generated summary of the key details. In summary, by producing readable abridgements from voluminous transcripts, this project aims to empower viewers to efficiently extract valuable information from online video news based on their interests and time constraints.

Motivation:

The driving force behind this project is the rapid growth of news outlets and consumer habits in the era of digital technology. The exponential expansion of

YouTube and other online video platforms has overtaken traditional summarizing methods as online video becomes the primary news source. Novel techniques that can extract meaningful information from lengthy video transcripts with different channel-specific terminology and structural variations are desperately needed. The main goal is to use the most recent advancements in neural natural language processing techniques to create reliable models specifically designed for summarizing streaming news items. More specifically, this entails leveraging new transformer architectures that have established new standards for text summarizing tasks, such as BART and T5. The research attempts to utilize these models' characteristics, such as cross-task adaptability and bidirectional context modelling.

Significance:

This idea has many more potential effects than just improving the way people consume news. The application of machine learning and language processing techniques for highly summarized access is being advanced, expanding the potential of these cutting-edge technologies. The developed methods might be applied to other domains, such as making intelligent news content recommendations or meeting the demands of people with impairments in terms of accessibility. More generally, the skills developed are in line with the overarching objective of using technological advancement to change media environments and promote inclusive news consumption through user-centered designs that consider a range of preferences, skill levels, and usage scenarios. Beyond immediate uses, this research advances artificial intelligence in a way that makes digital experiences naturally responsive to user requirements, advancing the development of digital ecosystems that are more widely accessible. By responding to the needs of change. It contributes to broader advancements in fair access to digital spaces by utilizing intelligent summarizing to meet the demands of a changing news environment.

Model Experiment:

Data Collection:

Data Sources:

The primary training data source for this project remains the “Newspaper Text Summarization - CNN/Daily Mail” dataset sourced from Kaggle. Comprising of news articles from the two major online publications CNN and DailyMail, it provides a rich collection of authentic news reports on diverse topics including political events, business, sports, culture, and entertainment. Leveraging this comprehensive dataset allows for robust training of both the BART and T5 neural summarization models underpinning this work by exposing them to a substantial vocabulary and writing styles reflecting real-world news landscapes. This equips the models with generalized capabilities for summarizing unseen YouTube news video transcripts on the full spectrum of current news topics and events through transfer learning. Instead of limiting insights to news from the training distribution itself, models can develop holistic news summarization expertise. This data sourcing strategy aids the end goal of seamless, insightful summarization—reflecting the key details accurately for widely variant YouTube news content experienced by the consumers.

Preprocessing:

Data preprocessing is a crucial step in preparing information for the News Article Summarization model designed for YouTube news transcripts. This process entails extracting pertinent details, including article text and highlights, and ensuring that the data aligns with the specific requirements of the model. The goal is to streamline the input data, making it optimal for the summarization model to effectively analyze and generate concise summaries. This involves cleaning and

organizing the text, addressing any inconsistencies, and structuring the information in a format that enhances the model's ability to generate accurate and coherent news summaries for YouTube transcripts.

Model Training:

Dataset Details:

For training, a diverse dataset of news transcripts gathered from YouTube and Kaggle are utilized. This broad range of sources enables both BART and T5 models to learn and generate summaries for various video content.

Hyperparameters:

The BART and T5 models' hyperparameters are adjusted to better fit the distinct qualities of traditional news articles and YouTube news transcripts. To achieve optimal performance, this entails optimising parameters like as learning rates, batch sizes, and training epochs.

Training Challenges:

Over combining datasets makes it more difficult to handle linguistic style differences and adapt to various domains. Strategies are used to get over these obstacles, resulting in a more robust and adaptable news item summarization model.

Model Architecture:

The selected model architecture makes use of the encoder-decoder structure of the BART model, which is ideal for processing both traditional news articles and YouTube news transcripts. With the help of this architecture, crucial information from input sequences is effectively extracted, enabling the decoder to provide concise, clear summaries.

Key Components:

Encoder-Decoder Structure:

Both BART and T5 models leverage their encoder-decoder architectures to effectively capture contextual relationships within the text.

Attention Mechanisms and Positional Embeddings:

Both models leverage attention mechanisms to focus on crucial parts of the input sequence during processing. This focus on relevant details enhances their ability to capture nuanced information and context within diverse news sources. Additionally, positional embeddings are employed to maintain the sequential order of information, ensuring contextually accurate summarization.

Training Strategy:

Transfer learning from pre-trained embeddings is utilized for both models. This facilitates a more efficient training process by leveraging existing linguistic knowledge from pre-trained resources. This pre-trained knowledge allows the models to adapt to the unique linguistic nuances present in various news formats.

BART Model:

BART, a state-of-the-art neural network architecture, employs an encoder-decoder structure with a unique training approach. Its bidirectional training enables the

model to capture context from both preceding and succeeding words, while its auto-regressive training facilitates the generation of coherent text. Additionally, BART utilizes a denoising objective during training, further enhancing its ability to capture essential information. This combination of features positions BART as a powerful and adaptable model for various NLP tasks.

T5 Model:

T5, a groundbreaking text-to-text transfer transformer, offers a unified framework for diverse NLP tasks. It transforms tasks like summarization and translation into a common text-to-text format, simplifying model development and training. T5 employs an encoder-decoder structure similar to other transformer models, but its unique text-to-text framework sets it apart. Pre-trained on a massive corpus, T5 gains a generalized understanding of language, which it then leverages when performing specific tasks through fine-tuning with task-specific prompts. This flexible approach allows T5 to excel across a broad range of NLP challenges.

Model Evaluation and Results Interpretation:

The project uses the Rouge score (Rouge-1, Rouge-2, and Rouge-L) to evaluate summarization quality. These metrics measure the overlap between generated summaries and reference summaries, providing insights into the models' performance. Precision, recall, and F1 scores derived from Rouge scores offer further nuanced perspectives. These metrics guide the refinement and optimization of the models, ensuring consistent production of accurate and coherent summaries across diverse input formats.

Flask App Integration:

Web Application:

A Flask web application has been developed to provide users with an easy and intuitive way to access concise summaries of YouTube news articles. This application features a user-friendly interface that allows users to input YouTube news links and receive summaries generated by two powerful models: BART and T5.

Here's a breakdown of the interface:

1. Input:

A clear and concise input field labeled "YouTube News Link" allows users to easily paste the URL of the desired news article.

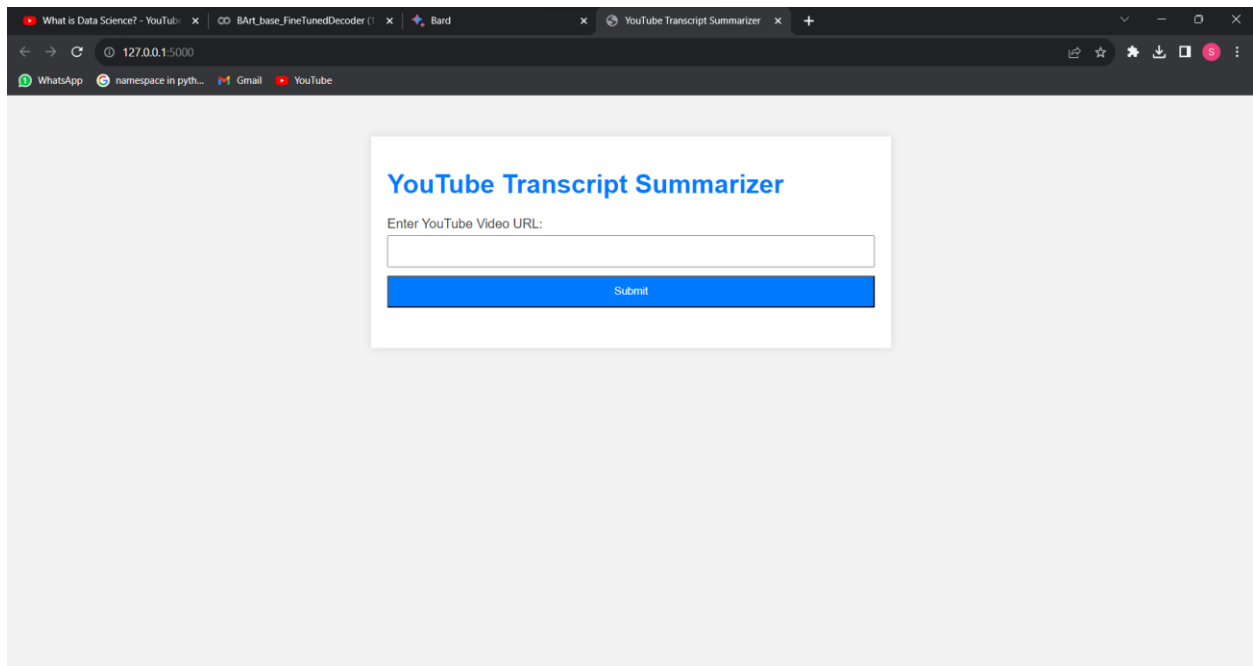
2. Summary Generation:

A prominent "Generate Summary" button triggers the process of fetching the news article content, extracting relevant information, and generating a concise summary using the model.

3. Output:

The generated summary is displayed prominently below the input field.

Below is the interface for the App look like:



Summary:

Key Findings:

The BART model has shown remarkable success in summarizing YouTube news video transcripts and traditional news articles. Its ability to adapt to diverse content formats and effectively handle domain-specific challenges highlights its versatility and potential for real-world applications.

BART's Strengths:

Adaptability: BART excels at handling both formal and informal language, making it well-suited for summarizing both traditional news articles and the conversational style of YouTube transcripts.

Performance: BART consistently produces high-quality summaries, achieving superior results on key evaluation metrics like Rouge scores. This demonstrates its ability to accurately capture essential information from diverse news sources.

Efficiency: BART's smaller model size compared to other options like T5 translates to faster training and deployment, making it more resource-efficient and scalable for real-world applications.

Future Work:

- **Advanced Architectures:** Exploring more advanced transformer architectures like T6 or Megatron-Turing NLG could potentially further enhance BART's performance and capabilities.
- **News-Specific Pre-training:** Experimenting with diverse pre-training strategies specifically tailored to news content could lead to improved knowledge and context understanding for BART.
- **Personalized Summarization:** Incorporating user feedback through interactive interfaces or preference polls could enable BART to personalize summaries to individual user preferences.

Conclusion:

The BART model's effectiveness in summarizing diverse news content makes it a promising tool for enhancing information access and comprehension. While further research and development can refine its capabilities, BART's current performance highlights its potential to revolutionize the way we interact with news and information.

[Git Repository Link](#)

References:

- <https://huggingface.co/docs/transformers/tasks/summarization#load-billsum-dataset>
- <https://medium.com/analytics-vidhya/text-summarization-using-nlp-3e85ad0c634>

