Paper ID: 807

ROUGE Score Analysis & Performance Evaluation Between Google T5 & SpaCy for YouTube News Video Summarization

Authors:

Nurshat Fateh Ali, Jamal Uddin Tanvin, MD Rifat Islam, Jubair Ahmed, M. Akhtaruzzaman

E-mail: nurshatfateh@gmail.com, jamaluddintanvin@outlook.com, rifat010bushra@gmail.com, jubairahmed1678@gmail.com, akhter900@cse.mist.ac.bd

Department of Computer Science and Engineering, Military Institute of Science and Technology, Mirpur Cantonment, Dhaka-1216, Bangladesh

Table of contents

01

Introduction

Introduction

The study aims to compare the performance between Google T5 and SpaCy for YouTube video textual summarization based on the ROUGE score.

The study includes the implementation of two distinct techniques: abstractive summarization using Google T5 and extractive summarization using SpaCy.

02

Objectives

Objectives

1

To implement a procedure for abstractive YouTube news video textual summarization with Google T5.

2

To implement a procedure for extractive YouTube news video textual summarization with SpaCy.

3

To evaluate and compare two approaches of YouTube news video textual summarization using ROUGE scores.

03

Motivation

Motivation

Long News Videos	Acceptability
The viewer may not have enough time to watch the entire news video. Summarization methods can help in this circumstances to consume news efficiently.	Only a comprehensive analysis and comparison between multiple approaches can help finding the best method for higher user acceptability.

04

Methodology

Methodology

Google T5

Google T5 is a text-to-text transfer transformer that is designed to perform advanced natural language processing (NLP) tasks.

SpaCy

Spacy is a free, open-source library for advanced Natural Language Processing (NLP) in Python.

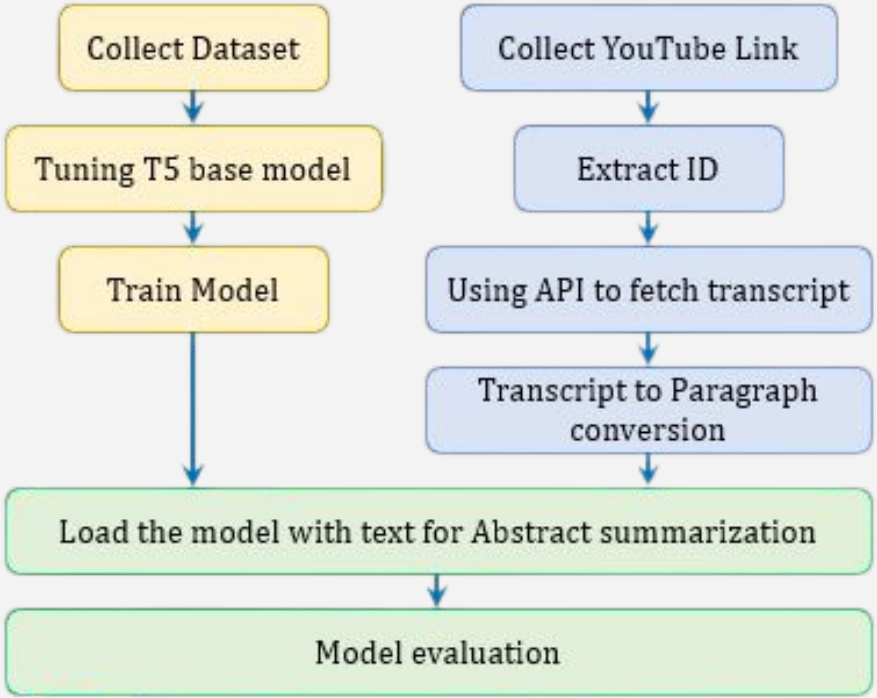
Dataset

CNN/DailyMail dataset from the Hugging face

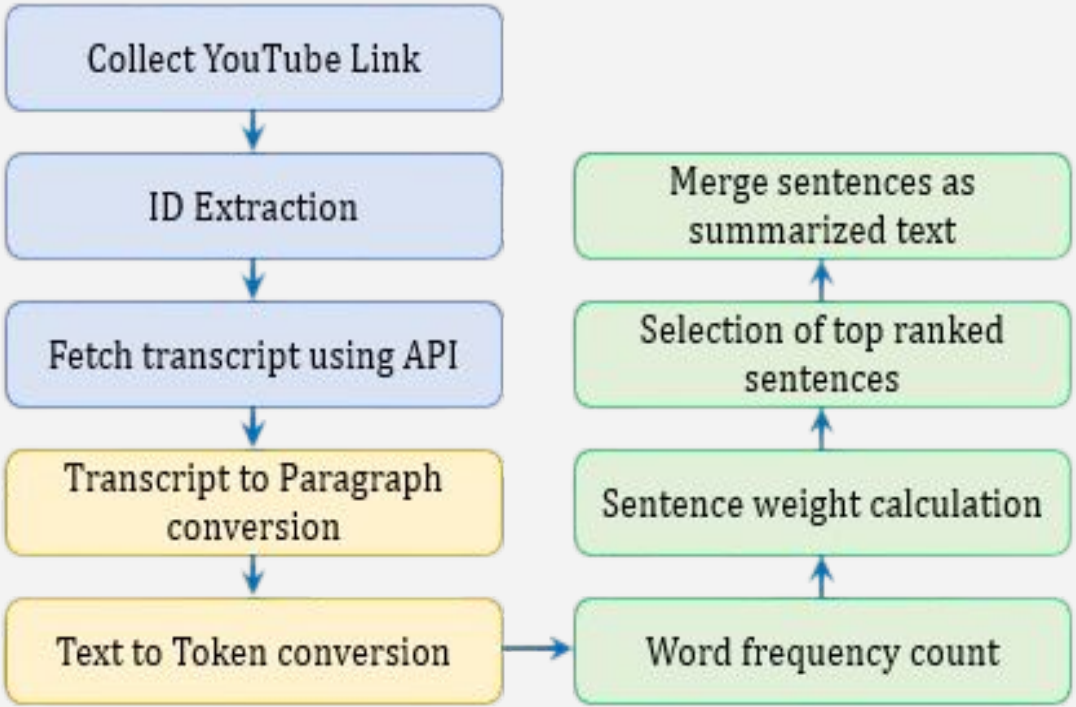
The dataset contains over 300k unique news articles with corresponding summaries.

article string	**highlights** string	**id** string
LONDON, England (Reuters) -- Harry Potter star Daniel Radcliffe gains…	Harry Potter star Daniel Radcliffe gets £20M fortune…	42c027e4ff9730fbb3de84c1af0d2c506e41c3e4
Editor's note: In our Behind the Scenes series, CNN correspondents…	Mentally ill inmates in Miami are housed on the "forgotten…	ee8871b15c50d0db17b0179a6d2beab35065f1e9
MINNEAPOLIS, Minnesota (CNN) -- Drivers who were on the Minneapoli…	NEW: "I thought I was going to die," driver says . Man…	06352019a19ae31e527f37f7571c6dd7f0c5da37

Abstractive Summarization



Extractive Summarization



Evaluation using ROUGE

Recall-Oriented Understudy for Gisting Evaluation

ROUGE-1	Refers to the overlap of unigrams between the system summary and reference summary.
ROUGE-2	Refers to the overlap of bigrams between the system and reference summaries.
ROUGE-L	Measures the longest matching sequence of words using LCS.
ROUGE-Lsum	A variant of ROUGE-L that computes the average LCS score between the system and reference summaries.

ROUGE-N

ROUGE-N is an n-gram recall between a candidate summary and a set of reference summaries.

ROUGE-N is computed as follows:

$$\text{ROUGE} - \text{N} = \frac{\sum_{S \in \{\text{Ref Summaries}\}} \sum_{n-gram \in S} Count_{match}(n-gram)}{\sum_{S \in \{\text{Ref Summaries}\}} \sum_{n-gram \in S} Count(n-gram)}$$

Here, n is the length of the n-gram & $Count_{match}$(n-gram) is the maximum number of n-gram co-occurring in a candidate summary and a set of reference summaries. Count(n-gram) is the number of n-grams in a sentence.

- ROUGE-1 (ROUGE-N for n=1) : overlap of unigrams
- ROUGE-2 (ROUGE-N for n=2) : overlap of bigrams

ROUGE-L

Measures the longest matching sequence of words using LCS.

ROUGE-L is simply using LCS-based F-Score measure to estimate the similarity between two summaries X of length m and Y of length n, assuming X is a reference summary sentence and Y is a candidate summary sentence, as follows:

$$R_{lcs} = \frac{LCS(X,Y)}{m}$$

$$P_{lcs} = \frac{LCS(X,Y)}{n}$$

$$F_{lcs} = \frac{(1+\beta^2)R_{lcs}P_{lcs}}{R_{lcs}+\beta^2 P_{lcs}}$$

Here LCS(X, Y) is the length of the longest common subsequence of X and Y , and $\beta = P_{lcs}/R_{lcs}$. **F_{lcs}** is considered as the ROUGE-L score.

ROUGE-Lsum

A variant of ROUGE-L that computes the average LCS score between the system and reference summaries.

$$R_{lcs} = \frac{\sum_{i=1}^{u} LCS_{\cup}(r_i, C)}{m}$$

$$P_{lcs} = \frac{\sum_{i=1}^{u} LCS_{\cup}(r_i, C)}{n}$$

$$F_{lcs} = \frac{(1+\beta^2) R_{lcs} P_{lcs}}{R_{lcs} + \beta^2 P_{lcs}}$$

Here, $LCS_{\cup}(r_i, C)$ is the LCS score of the union longest common subsequence between reference sentence r_i and candidate summary C and $\beta = P_{lcs}/R_{lcs}$.
F_{lcs} is considered as the ROUGE-Lsum score.

05

Results

Example Summary-1

YouTube link: https://www.youtube.com/watch?v= pHglSTvgbI

Video Title: The Russians leaving their country for Finland - BBC News

Provided Human Summary: Some Russian people are anxious to get out of the country because there has been a persistent rumour that President Vladimir Putin's government might soon introduce martial law to deal with demonstrations against the invasion of Ukraine. With flights to Europe halted, the only way out of the country is by car - crossing this border - or by train.

T5 Summary: People in finland want to give up old neutrality and join the western alliance. there are rumors that president putin will soon introduce martial law. for russians coming to finland is an escape from the dangers of life there.

SpaCy Summary 10%: For russians coming to finland is an escape from the dangers of life there, but for people on this side of the border there's real fear that the tensions within russia could boil over and engulf finland itself.

ROUGE Scores for Example Summary-1

ROUGE Scores	Google T5	SpaCy
ROUGE-1	0.327	0.265
ROUGE-2	0.083	0.021
ROUGE-L	0.286	0.122
ROUGE-Lsum	0.286	0.122

Example Summary-2

YouTube link: https://www.youtube.com/watch?v=6VZfeexme4c

Video Title: North Korean leader Kim Jong Un enters Russia to visit President Putin - BBC News

Provided Human Summary: North Korea's leader Kim Jong Un has crossed the border into Russia for a meeting with President Vladimir Putin. They are likely to discuss an arms deal as Russia faces a Ukrainian counter offensive, a US official said. South Korea's defence ministry confirmed that Mr Kim's armoured train entered Russia early on Tuesday morning.

T5 Summary: Kim Jong-un's train entered Russia on his way to meet with President Putin. Mr Kim is believed to be meeting to finalize an arms deal with Russia. North Korea has become increasingly isolated, according to reports.

SpaCy Summary 10%: It's also thought that Mr Kim might ask in return for these weapons, might ask Mr Putin to hand over sensitive weapons technology that he could then use to make breakthroughs in his own nuclear weapons program, though this isn't something that officials here are especially concerned about. it's taking place in Vladivostok and, according to the Kremlin, at the meeting with Kim Jong-un, will be taking place at the conclusion of that event.

ROUGE Scores for Example Summary-2

ROUGE Scores	Google T5	SpaCy
ROUGE-1	0.458	0.221
ROUGE-2	0.191	0.075
ROUGE-L	0.270	0.103
ROUGE-Lsum	0.271	0.103

The Overall ROUGE Scores for Google T5 and SpaCy

ROUGE Scores	Google T5	SpaCy
ROUGE-1	0.339	0.232
ROUGE-2	0.092	0.047
ROUGE-L	0.219	0.137
ROUGE-Lsum	0.218	0.136

06

Discussion

Key Findings

Google T5 outperforms SpaCy every single time.

The overall ROUGE-1 score for Google T5 is found to be 0.339.

The overall ROUGE-1 score for SpaCy is found to be 0.232.

Highlights

Google T5 summarization shows better unigram overlapping with the summaries generated by humans.

Google T5 presents better results (10% higher) than SpaCy in terms of the ROUGE-1 score.

The potential of Google T5 in summarizing YouTube news videos is very promising.

07

Conclusion

It is observed...

Clearly

Abstractive summarization method using Google T5 outperforms extractive summarization method SpaCy in every single metric.

Also

Higher scores for Google T5 certifies the superiority of the transformer-based models for summarization.

07

Future Work

Future Work

Future work of this project can be focussed on enhancing the effectiveness and applicability of the summarization system by,

1. **Development of Browser Extension**
2. **Support of Multiple Languages**


Thank You!