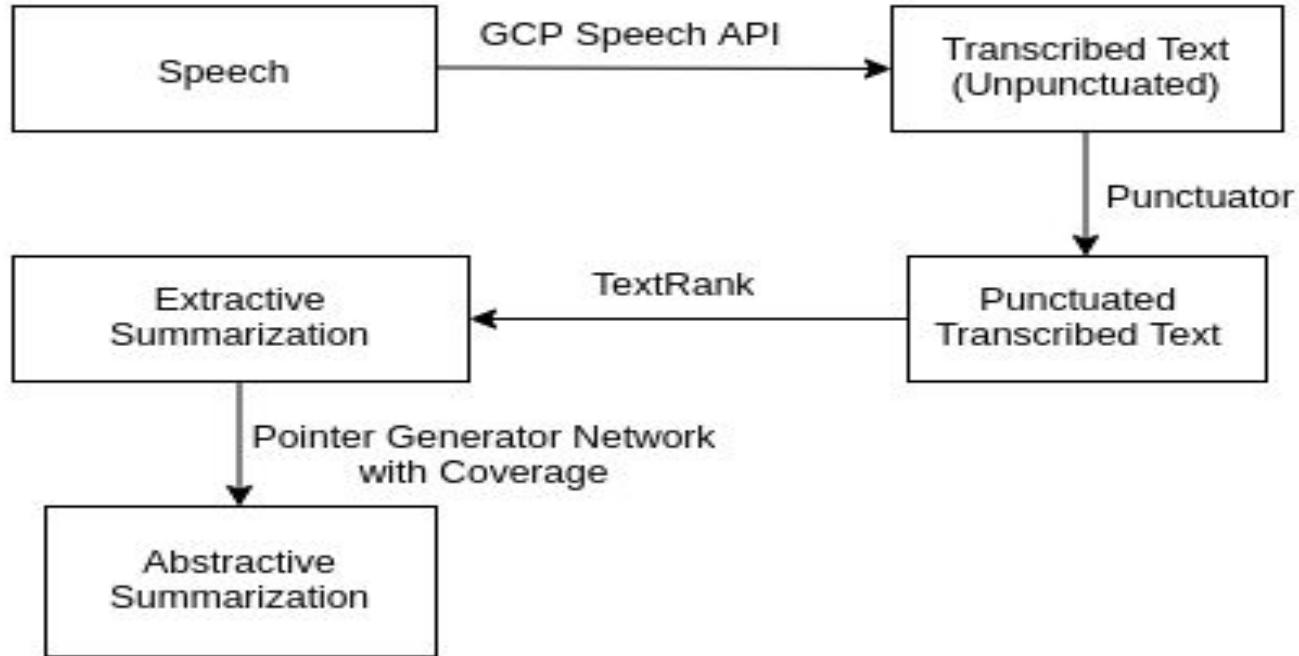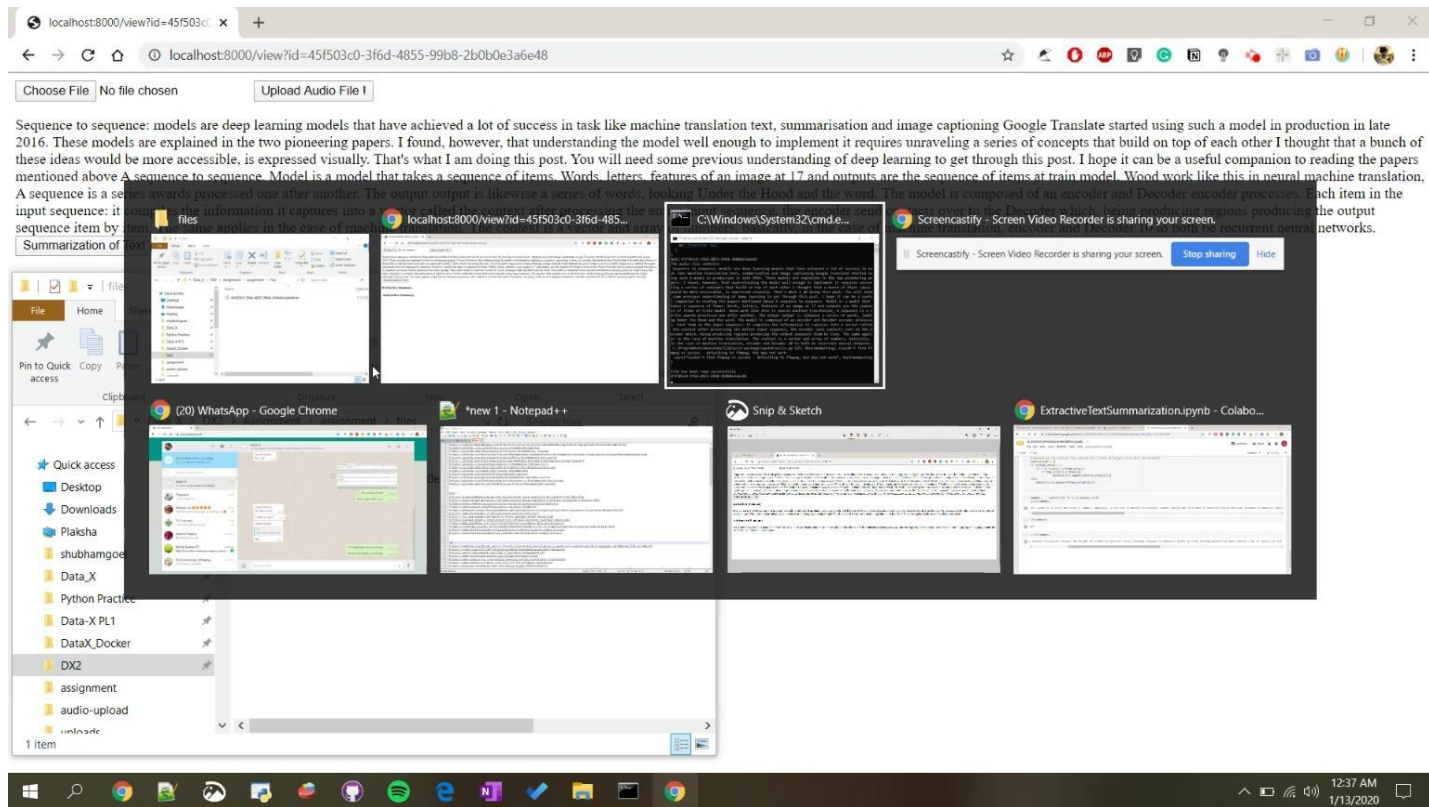# Speech to Text Summarization

Collaborators:
Karthika, Mahathi, Rashmi, Shikhar, Shubham Goel

Workflow

Automatic Speech Summarization System with text presentation
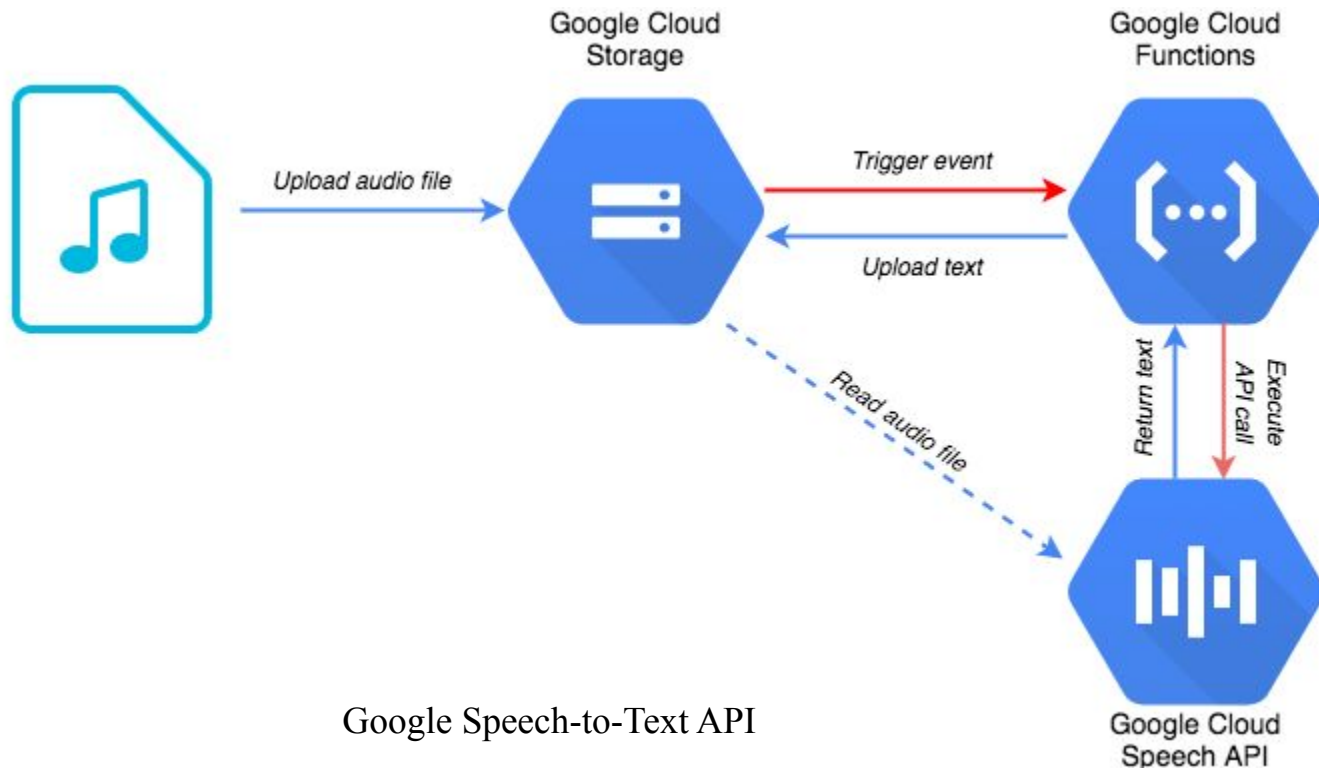
Demo

Datasets Used

We created a dataset for the summarization task as the current datasets weren't very suitable for our task.
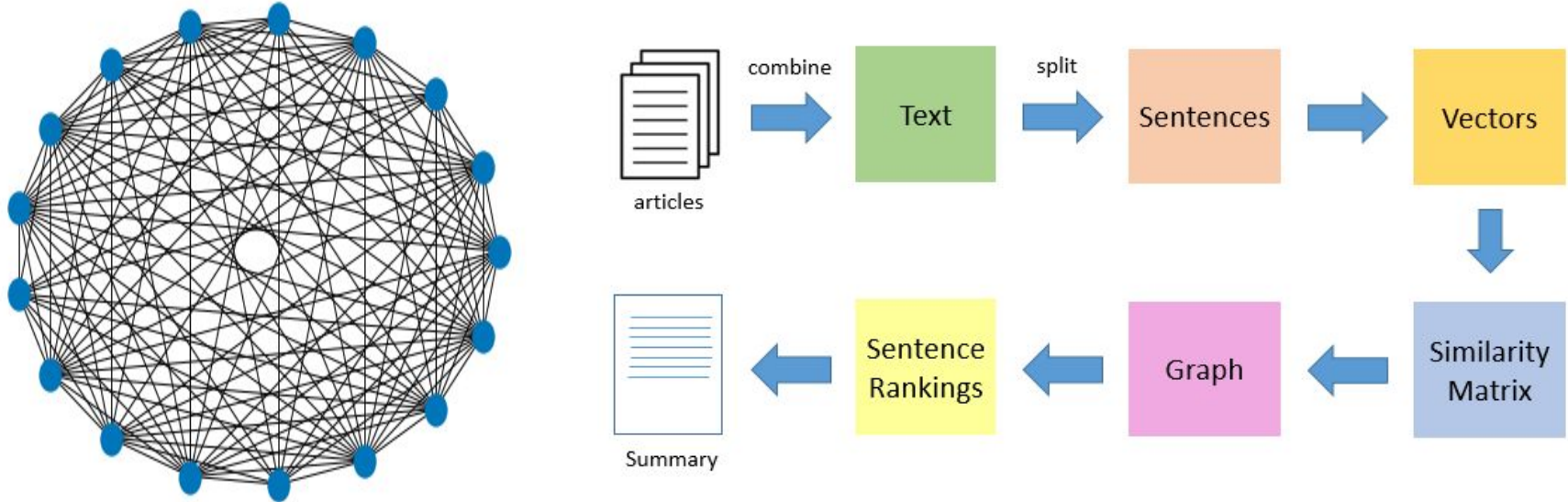
- The CNN/ Daily mail Dataset has very short summaries (in the form of bullet points).
- Inshorts has people manually summarizing news articles (in 60 words), hence makes more sense for our task.

	author	date	headlines	read_more	text	ctext
0	Chhavi Tyagi	03 Aug 2017,Thursday	Daman & Diu revokes mandatory Rakshabandhan in...	http://www.hindustantimes.com/india-news/raksh...	The Administration of Union Territory Daman an...	The Daman and Diu administration on Wednesday ...
1	Daisy Mowke	03 Aug 2017,Thursday	Malaika slams user who trolled her for 'divorc...	http://www.hindustantimes.com/bollywood/malaik...	Malaika Arora slammed an Instagram user who tr...	From her special numbers to TV? appearances, Bo...

Speech to Text
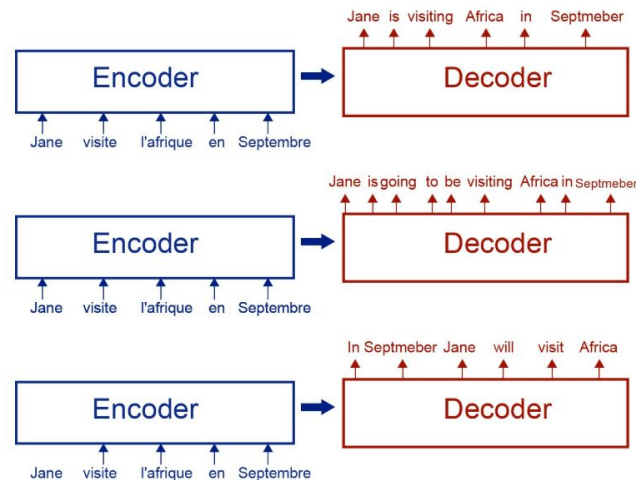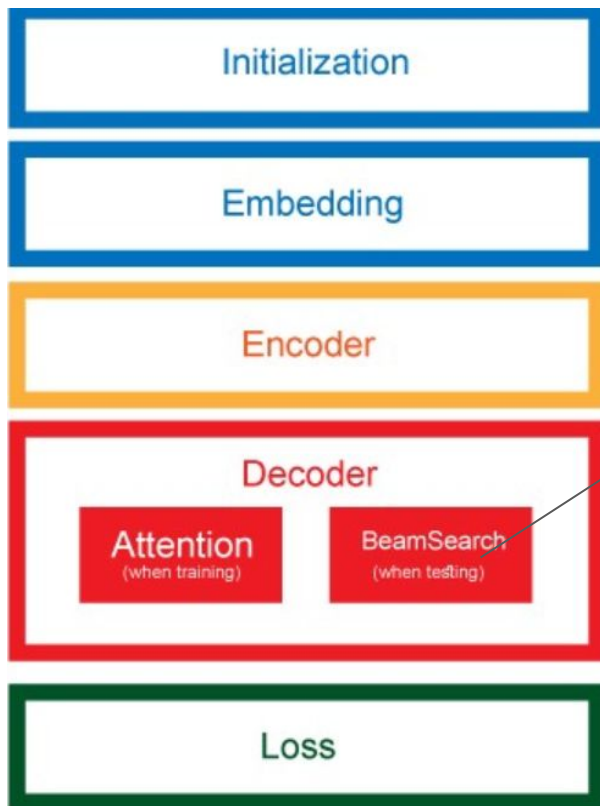

Google Speech-to-Text API

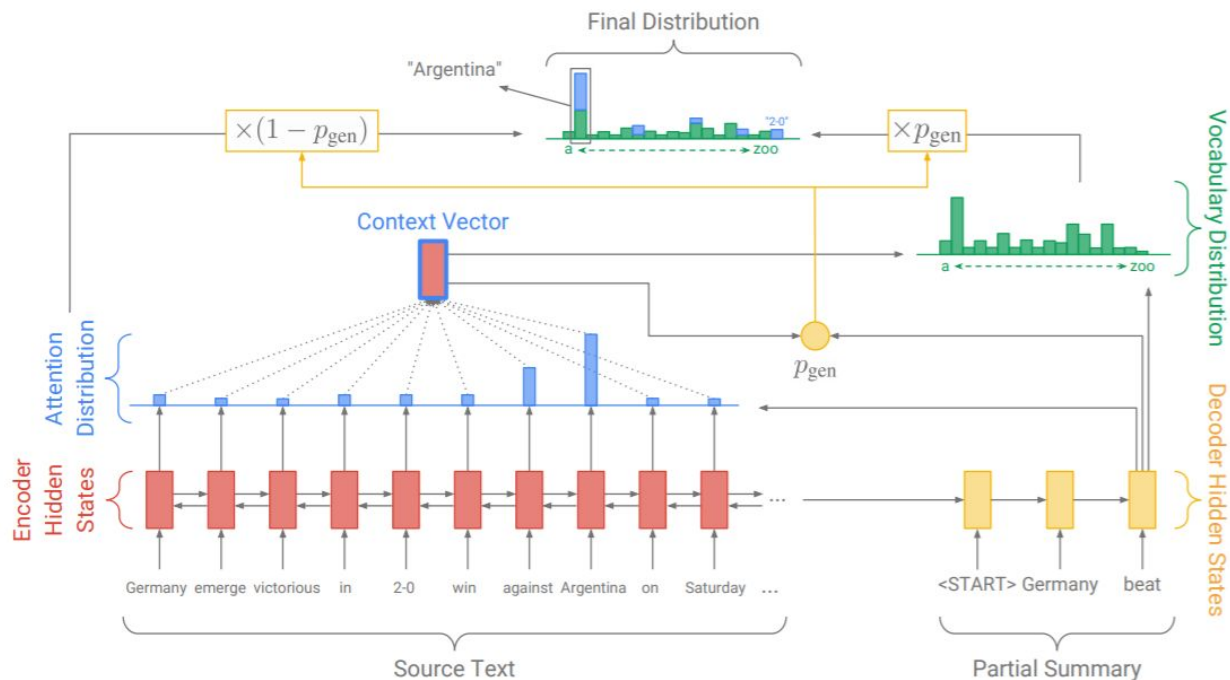# Extractive Summarization (TextRank)

Each node represents a sentence, while each edge represents words in common

Abstractive Summarization (Seq2Seq Attention)

Abstractive Summarization (Pointer Generator)

Integration with Front-end

Steps -

1. Make HTML UI to upload audio file and show summarised texts

2. Python Code to convert speech to text and save to text file

3. Run conversion code from server

4. Show text on Front-end

5. Show Summarise action button

5. Summarise (Extractive & Abstractive) above shown text in python and save to text file

6. Run summarisation code (Extractive & Abstractive) from server

7. Show summarized text (Extractive & Abstractive) on Front-end

Thank You