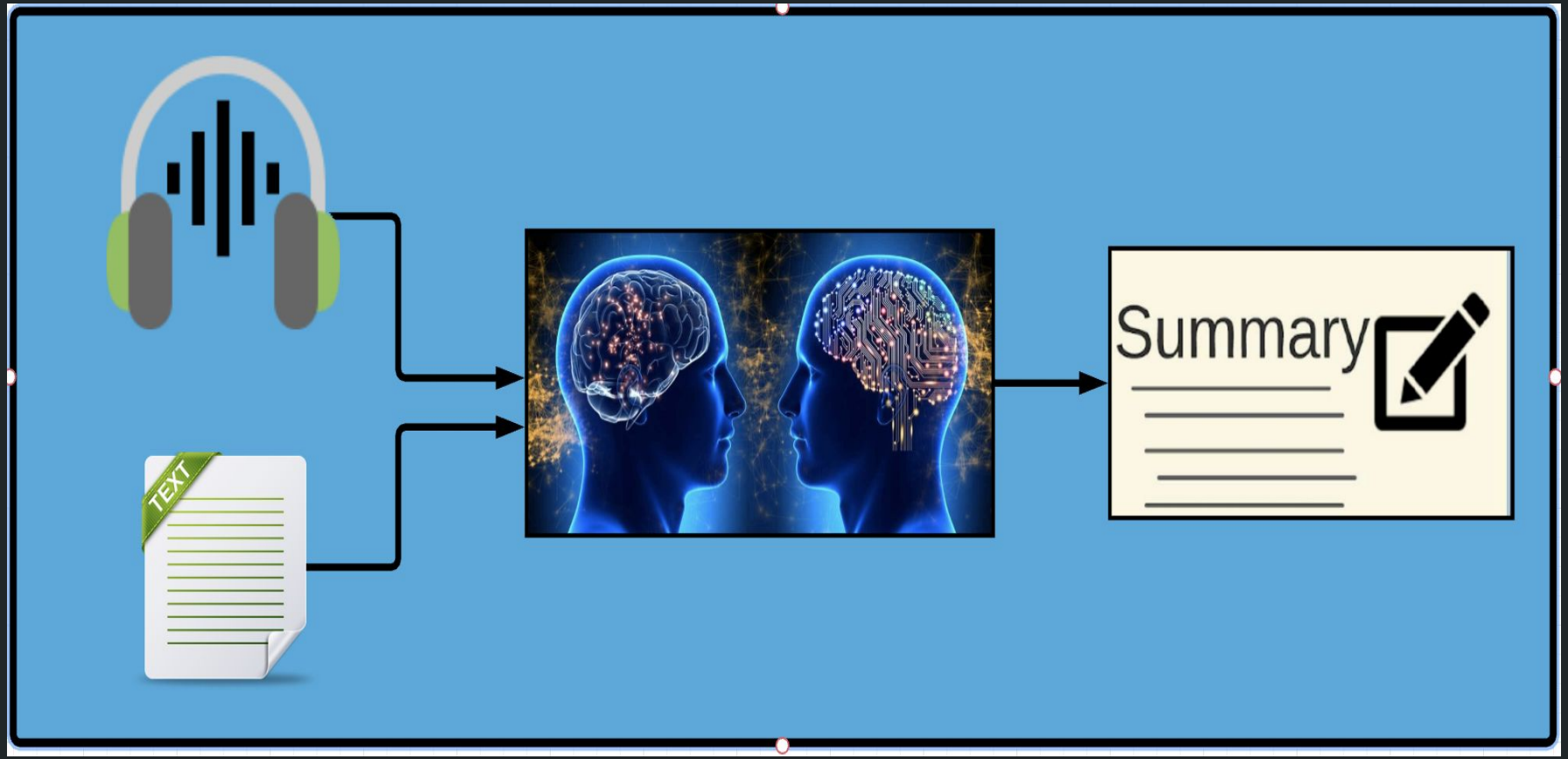
Text Summarization

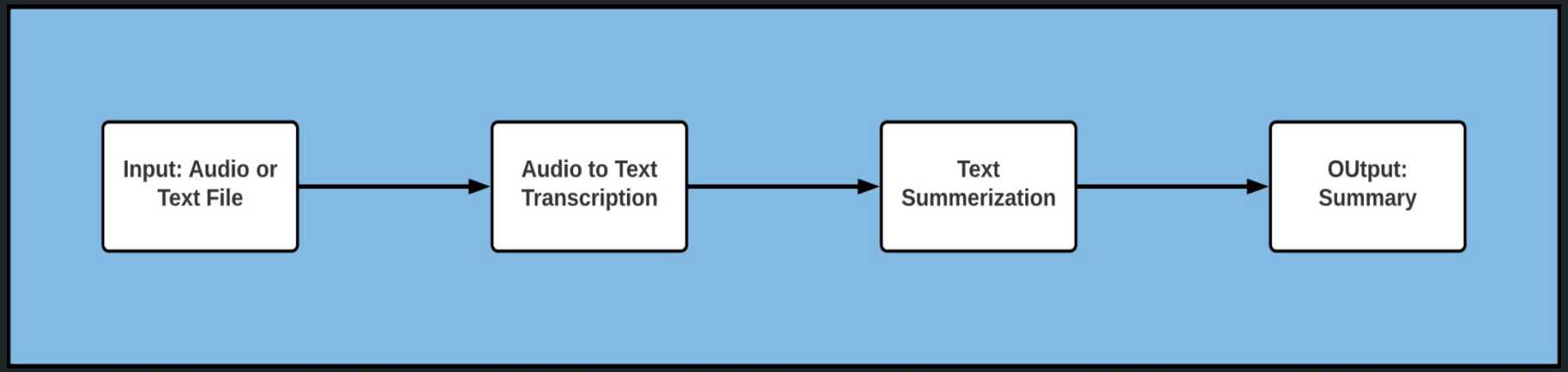
- Akanksha Rawat
- Karishma Kuria
- Rajvi Shah


TEXT
Summary

PROJECT OVERVIEW

- Generates succinct, accurate, easy to read, and comprehensive summaries of lengthy Text and Audio files.
- Different text summarization methods/model are implemented to get best condensed text summary.
- Reading a condensed summary can give anyone a sense of what audio or text files focuses on or help catch up on hours’ worth of content in minutes.

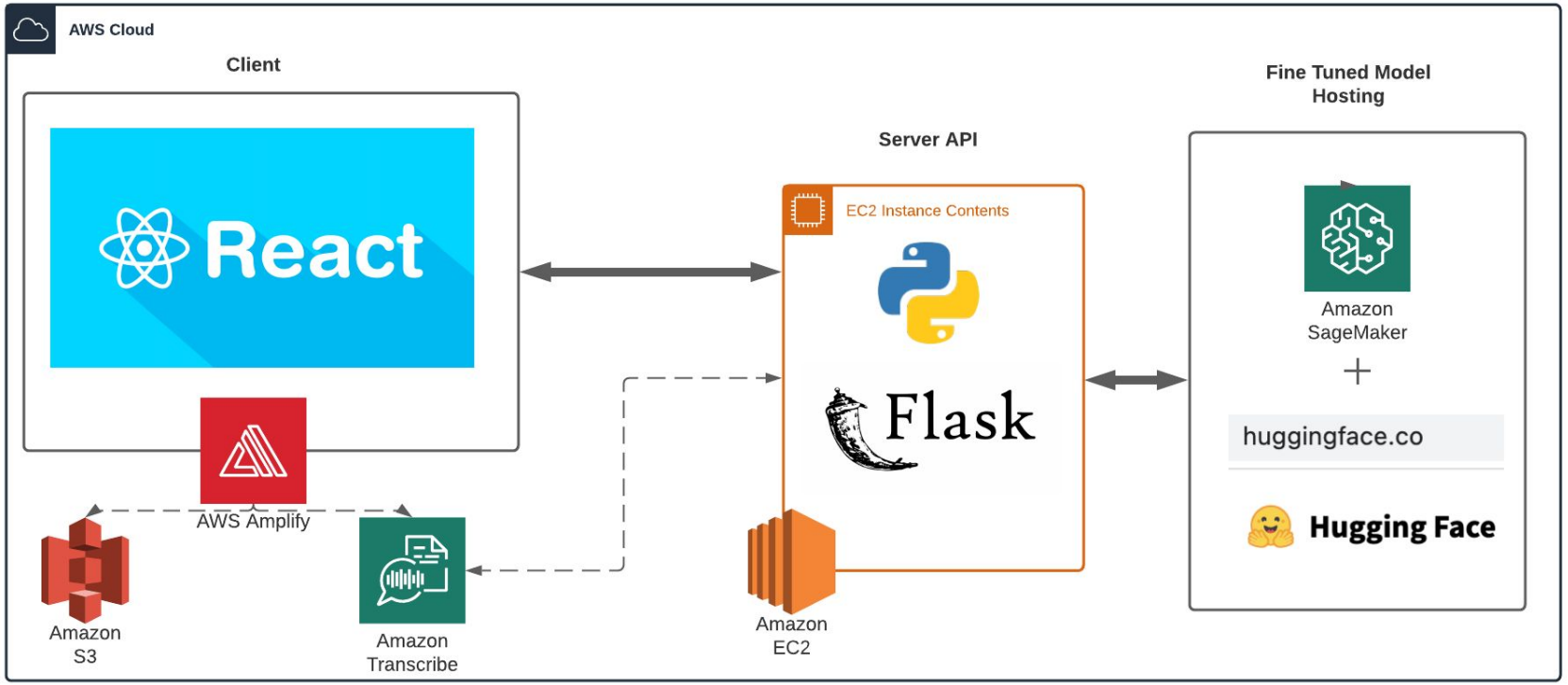
PIPELINE OF THE PROJECT



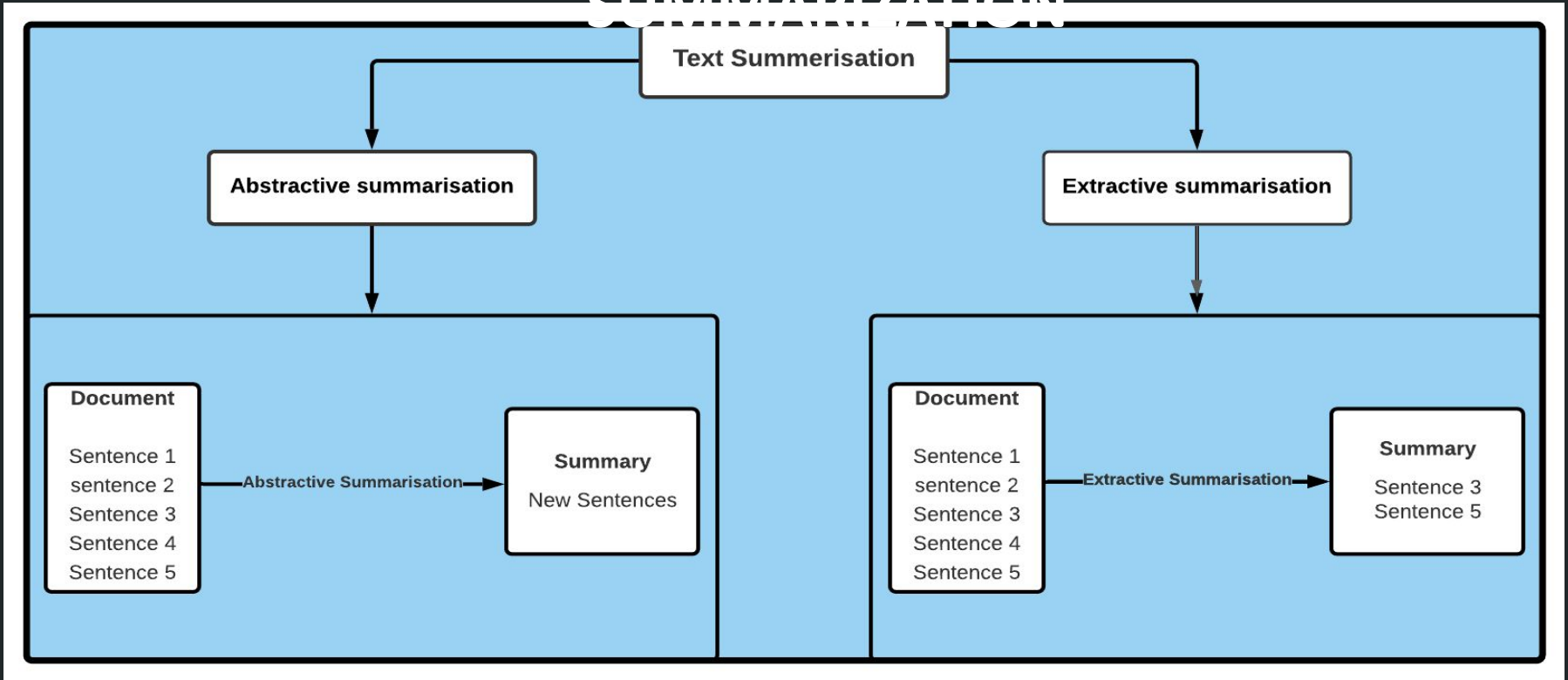
Two phases in the pipeline:

- Input- take input as audio or text file, convert Speech to text, display input as text.
- Output- Text-to-Text Summarization- Summary.

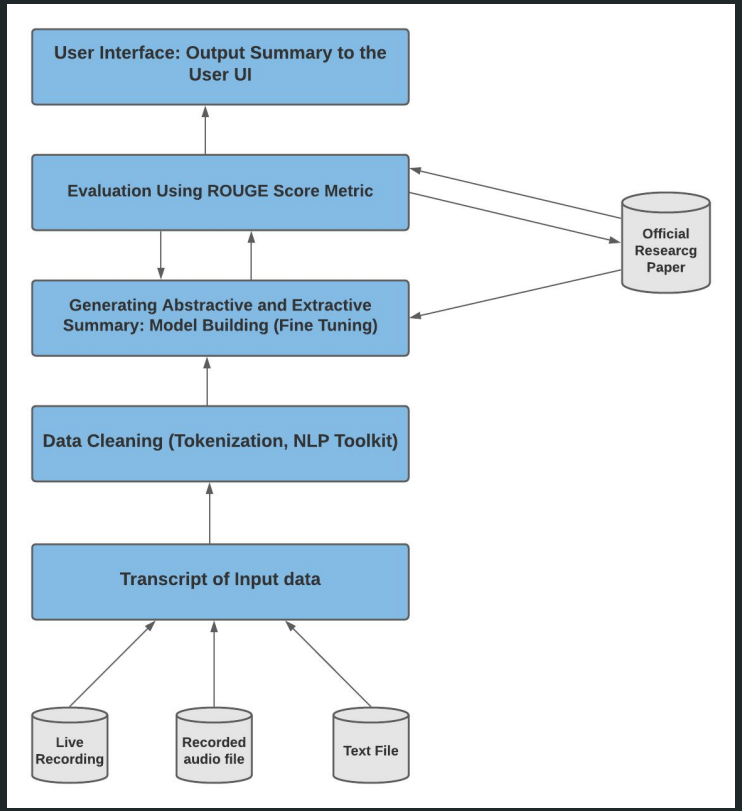
COMPONENTS


AWS Cloud
Client
React
AWS Amplify
Amazon S3
Amazon Transcribe
Server API
EC2 Instance Contents
Flask
Amazon EC2
Fine Tuned Model Hosting
Amazon SageMaker
+
huggingface.co
Hugging Face

METHODS USED FOR TEXT SUMMARIZATION



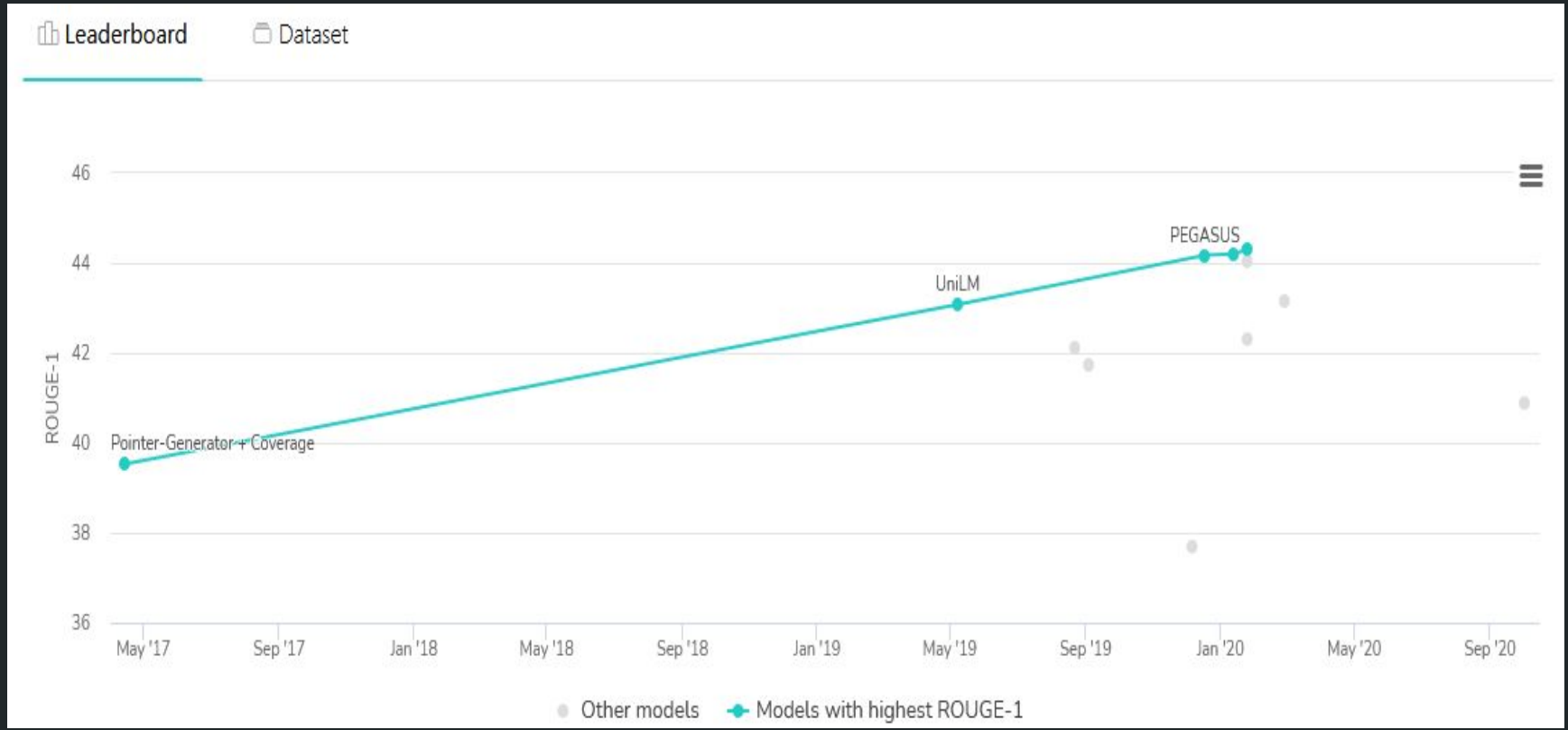
HIGH-LEVEL ARCHITECTURE DESIGN AND DATA FLOW DIAGRAM



WHY PEGASUS ?

- Google’s PEGASUS further improved the state-of-the-art (SOTA) results for abstractive summarization in particular with less resources.
- Uses Encoder-decoder model for sequence to sequence learning.
- Gap sentence Generation
- Masked Language Model
- ROUGE Score: to select important sentence.

LEADERBOARD FOR ABSTRACTIVE SUMMARIZATION



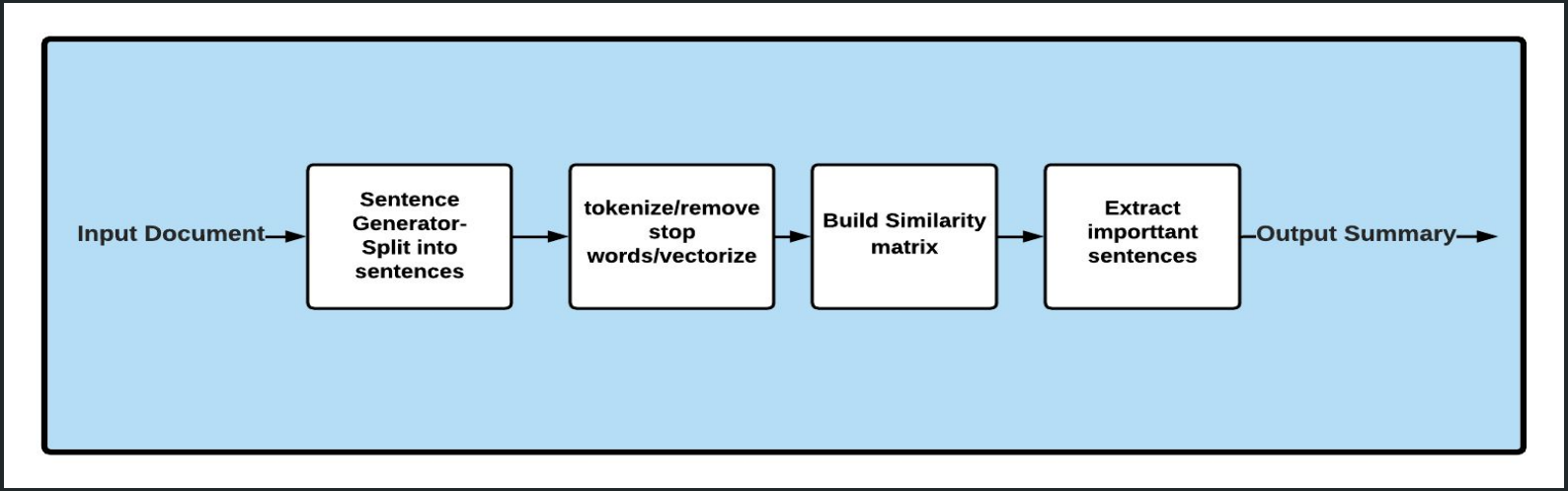
Implementation Insights

- Unsupervised Extractive Summarization

Input article → split into sentences → remove stop words → build a similarity matrix → generate rank based on matrix → pick top N sentences for summary.

Implementation Insights

- Unsupervised Extractive Summarization



Implementation Insights

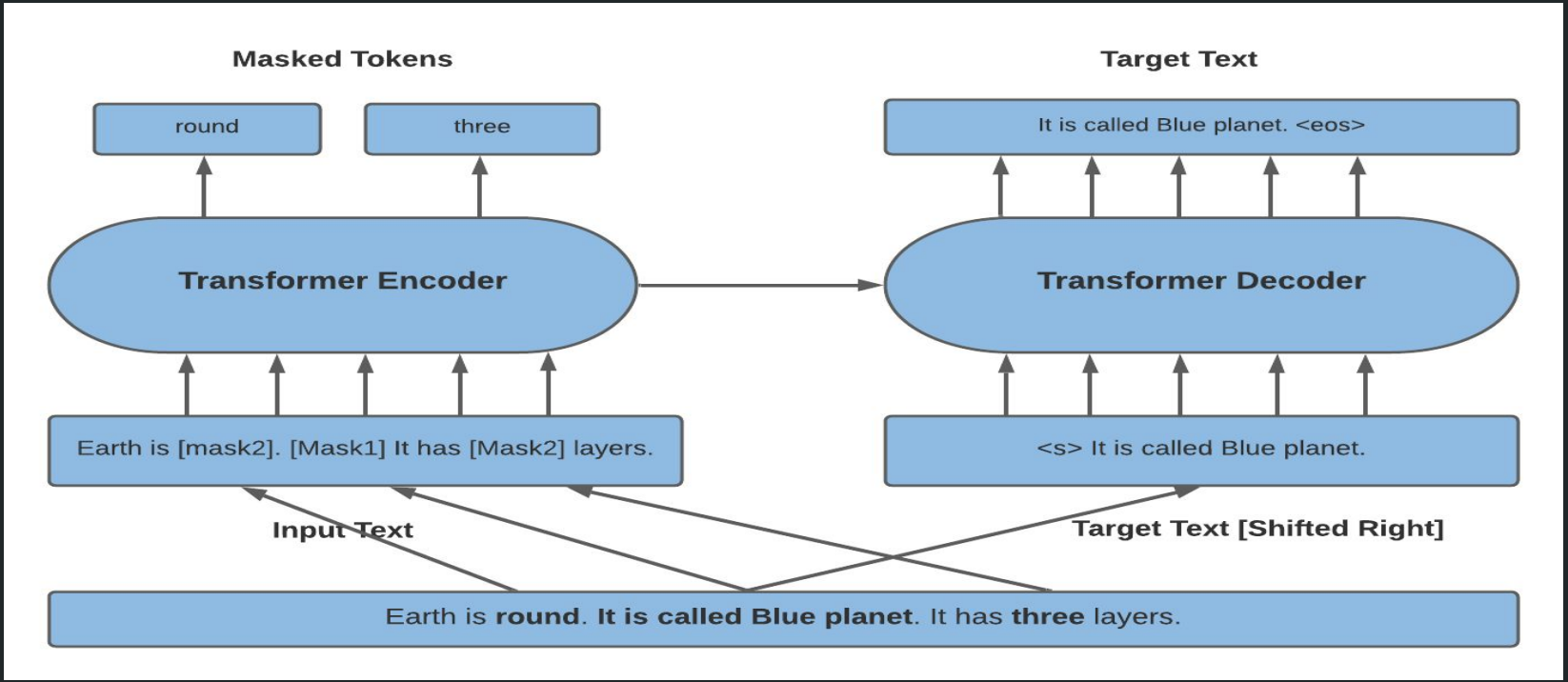
- Abstractive Summarization

  Fine-tuned CUSTOM_MODEL_NAME = 'arawat/pegasus-custom-xsum'

  Pretrained models-

1. google/pegasus-billsum
2. google/pegasus-cnn_dailymail
3. google/pegasus-multi_news

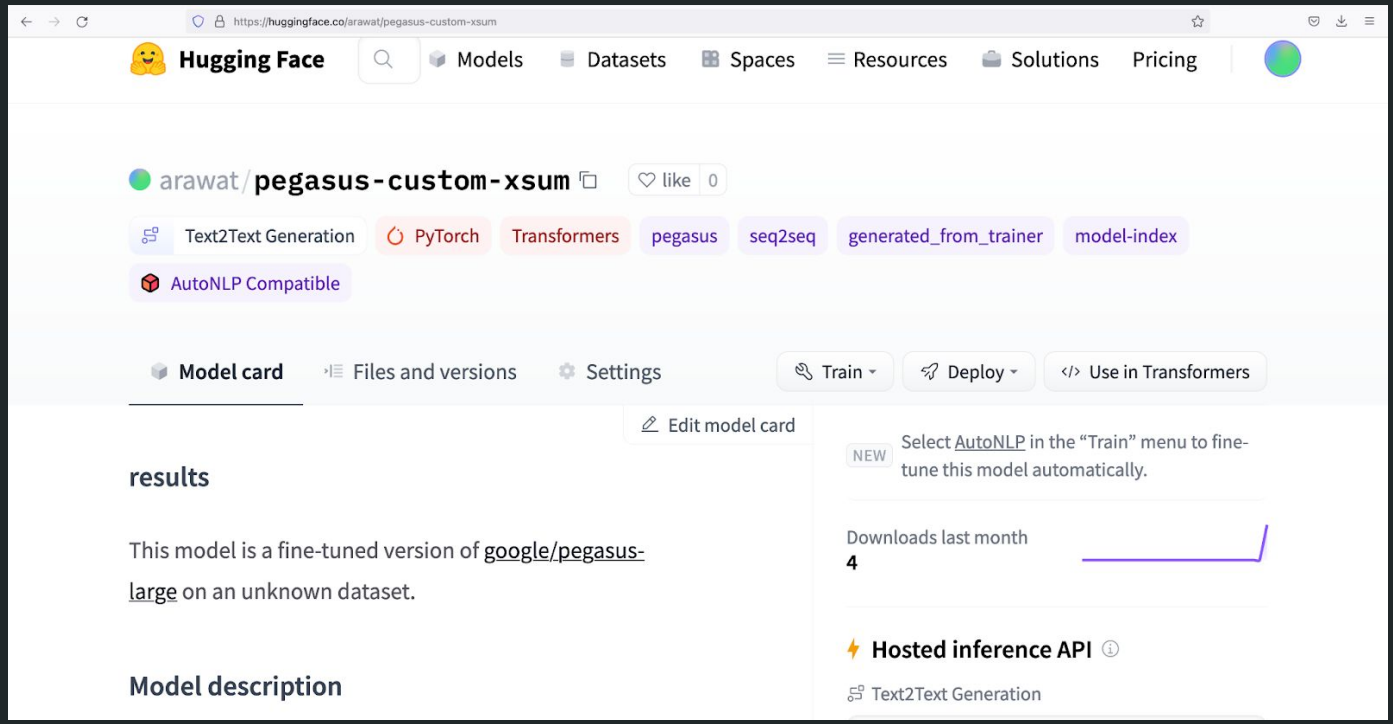
PEGASUS ARCHITECTURE



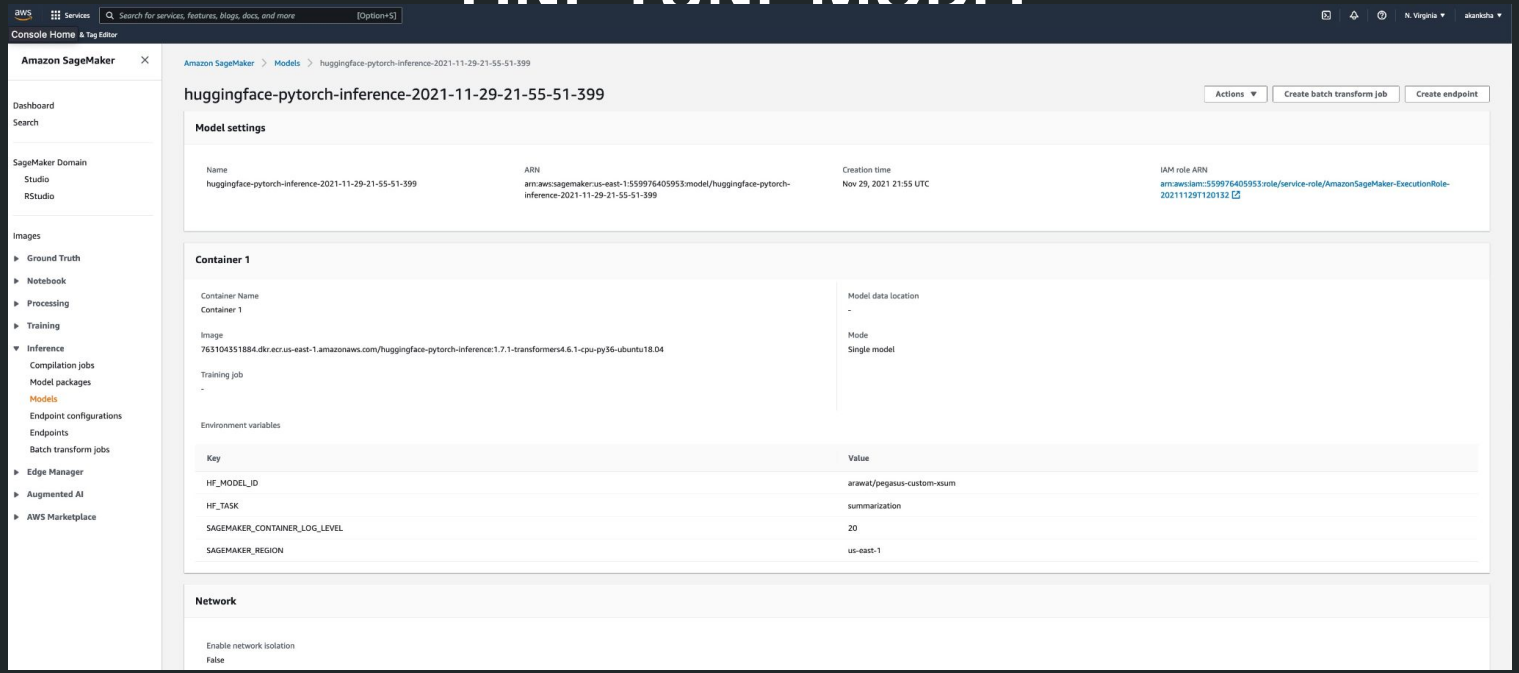
FINE-TUNE MODEL

- Custom- fine-tuned google/pegasus-large model on XSUM dataset
- Deployed on Huggingface Library/AWS Sagemaker

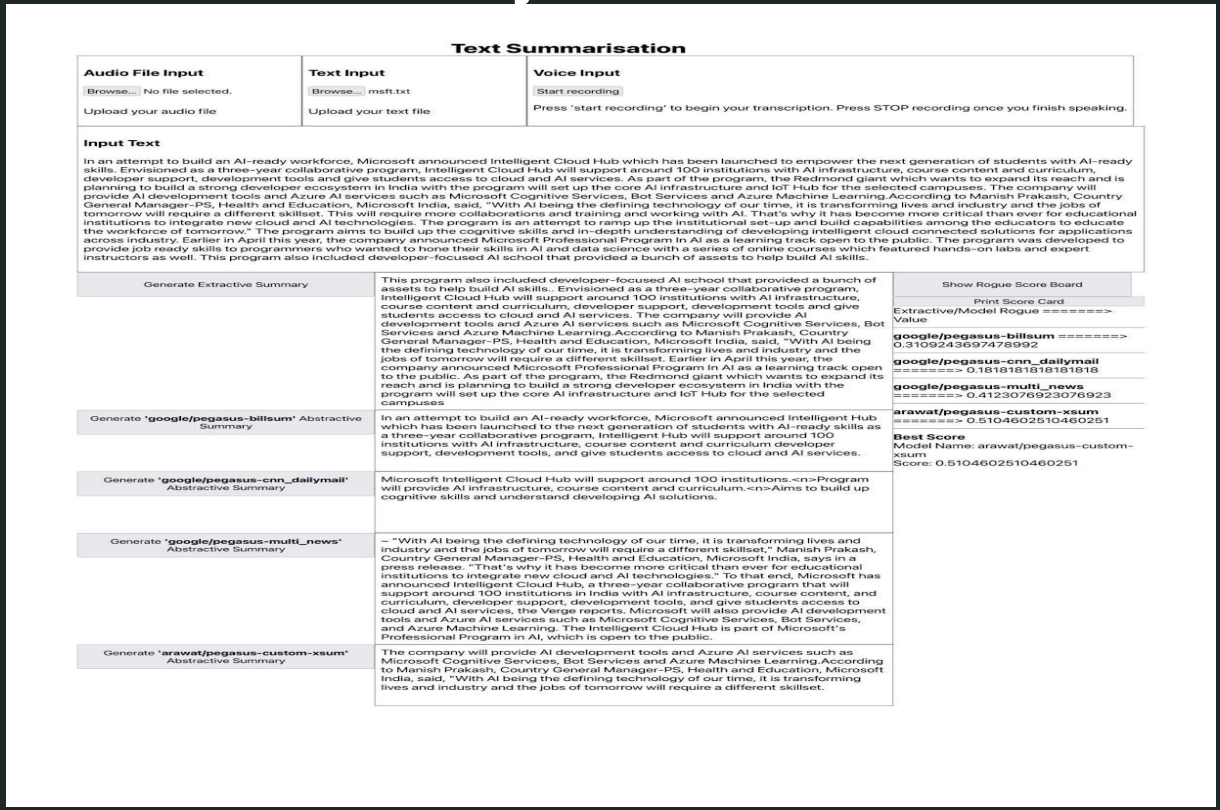
FINE-TUNE MODEL



FINE-TUNE MODEL



Project- Demo



Text Summarisation

Audio File Input

Browse... No file selected.

Upload your audio file

Text Input

Browse... msft.txt

Upload your text file

Voice Input

Start recording

Press 'start recording' to begin your transcription. Press STOP recording once you finish speaking.

Input Text

In an attempt to build an AI-ready workforce, Microsoft announced Intelligent Cloud Hub which has been launched to empower the next generation of students with AI-ready skills. Envisioned as a three-year collaborative program, Intelligent Cloud Hub will support around 100 institutions with AI infrastructure, course content and curriculum, developer support, development tools and give students access to cloud and AI services. As part of the program, the Redmond giant which wants to expand its reach and is planning to build a strong developer ecosystem in India with the program will set up the core AI infrastructure and IoT Hub for the selected campuses. The company will provide AI development tools and Azure AI services such as Microsoft Cognitive Services, Bot Services and Azure Machine Learning.According to Manish Prakash, Country General Manager-PS, Health and Education, Microsoft India, said, "With AI being the defining technology of our time, it is transforming lives and industry and the jobs of tomorrow will require a different skillset. This will require more collaborations and training and working with AI. That's why it has become more critical than ever for educational institutions to integrate new cloud and AI technologies. The program is an attempt to ramp up the institutional set-up and build capabilities among the educators to educate the workforce of tomorrow." The program aims to build up the cognitive skills and in-depth understanding of developing intelligent cloud connected solutions for applications across industry. Earlier in April this year, the company announced Microsoft Professional Program In AI as a learning track open to the public. The program was developed to provide job ready skills to programmers who wanted to hone their skills in AI and data science with a series of online courses which featured hands-on labs and expert instructors as well. This program also included developer-focused AI school that provided a bunch of assets to help build AI skills.

Generate Extractive Summary	This program also included developer-focused AI school that provided a bunch of assets to help build AI skills. Envisioned as a three-year collaborative program, Intelligent Cloud Hub will support around 100 institutions with AI infrastructure, course content and curriculum, developer support, development tools and give students access to cloud and AI services. The company will provide AI development tools and Azure AI services such as Microsoft Cognitive Services, Bot Services and Azure Machine Learning.According to Manish Prakash, Country General Manager-PS, Health and Education, Microsoft India, said, "With AI being the defining technology of our time, it is transforming lives and industry and the jobs of tomorrow will require a different skillset. Earlier in April this year, the company announced Microsoft Professional Program In AI as a learning track open to the public. As part of the program, the Redmond giant which wants to expand its reach and is planning to build a strong developer ecosystem in India with the program will set up the core AI infrastructure and IoT Hub for the selected campuses	Show Rogue Score Board Print Score Card Extractive/Model Rogue ========> Value **google/pegasus-billsum** ========> 0.3109243697478992 **google/pegasus-cnn_dailymail** ========> 0.18181818181818182 **google/pegasus-multi_news** ========> 0.4123076923076923 **arawat/pegasus-custom-xsum** ========> 0.5104602510460251 **Best Score** Model Name: arawat/pegasus-custom-xsum Score: 0.5104602510460251
Generate '**google/pegasus-billsum**' Abstractive Summary	In an attempt to build an AI-ready workforce, Microsoft announced Intelligent Hub which has been launched to the next generation of students with AI-ready skills as a three-year collaborative program, Intelligent Hub will support around 100 institutions with AI infrastructure, course content and curriculum developer support, development tools, and give students access to cloud and AI services.	
Generate '**google/pegasus-cnn_dailymail**' Abstractive Summary	Microsoft Intelligent Cloud Hub will support around 100 institutions.<n>Program will provide AI infrastructure, course content and curriculum.<n>Aims to build up cognitive skills and understand developing AI solutions.	
Generate '**google/pegasus-multi_news**' Abstractive Summary	– "With AI being the defining technology of our time, it is transforming lives and industry and the jobs of tomorrow will require a different skillset," Manish Prakash, Country General Manager-PS, Health and Education, Microsoft India, says in a press release. "That's why it has become more critical than ever for educational institutions to integrate new cloud and AI technologies." To that end, Microsoft has announced Intelligent Cloud Hub, a three-year collaborative program that will support around 100 institutions in India with AI infrastructure, course content, and curriculum, developer support, development tools, and give students access to cloud and AI services, the Verge reports. Microsoft will also provide AI development tools and Azure AI services such as Microsoft Cognitive Services, Bot Services, and Azure Machine Learning. The Intelligent Cloud Hub is part of Microsoft's Professional Program in AI, which is open to the public.	
Generate '**arawat/pegasus-custom-xsum**' Abstractive Summary	The company will provide AI development tools and Azure AI services such as Microsoft Cognitive Services, Bot Services and Azure Machine Learning.According to Manish Prakash, Country General Manager-PS, Health and Education, Microsoft India, said, "With AI being the defining technology of our time, it is transforming lives and industry and the jobs of tomorrow will require a different skillset.	

Any Questions...

Thank you!