

# Project: Speech/Text to Text Summarization

Project Team Name- Data Decoders

Project Team Members-

Akanksha Rawat, Karishma Kuria

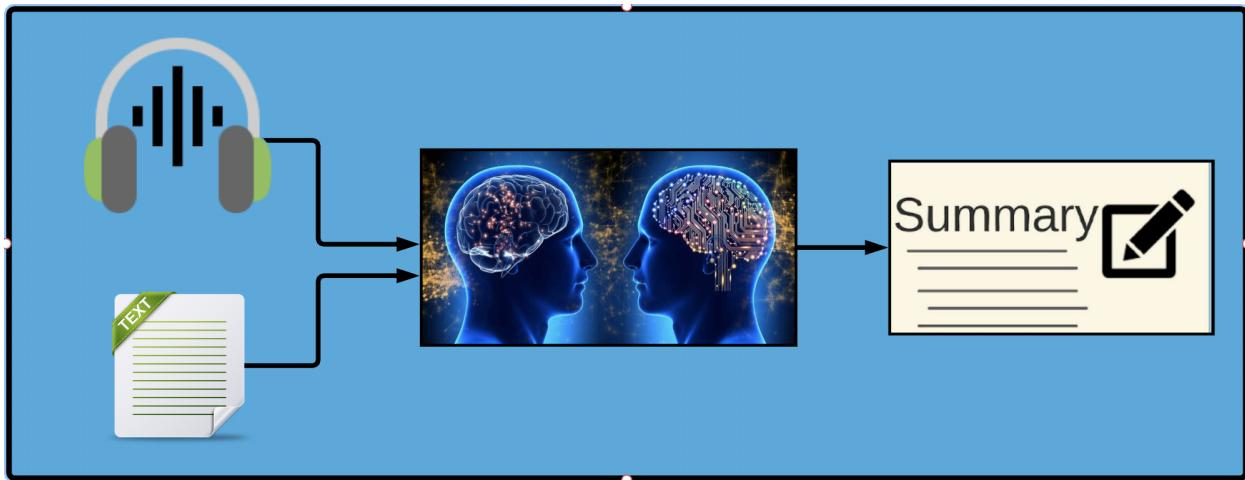
Rajvi Shah

CMPE 256 - Fall, 2021

12/17/2021



## **Abstract:**



**Description:** It intends to provide concise, accurate, easy to read, and comprehensive summaries of lengthy audio files or text files that one can read within a few minutes.

Reading a condensed summary of details from a long audio file/text file can give anyone a sense of what file focuses on or help catch up on hours' worth of content in minutes.

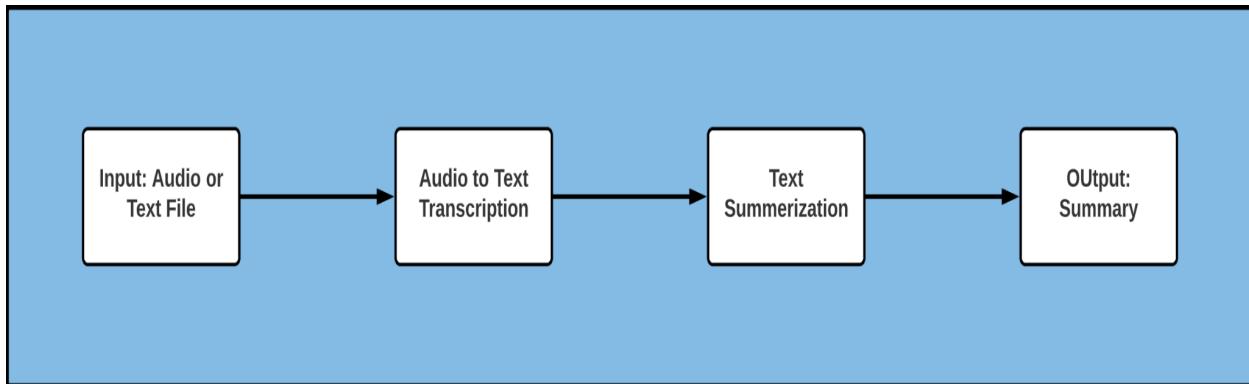
Due to digital media, the amount of online content available today is massive in today's era. Books, articles, blogs, and audios like podcasts, news broadcasting on radio channels, speeches, etc. All the audio files data obtained from these sources are not a valuable means of gathering information every time. It is always effective if the data is obtained from the condensed content i.e., focusing more on essential points than the entire content.

This application can be intensively used in summarizing business meetings. It serves in more efficient progress tracking in project meetings and facilitation of learning using online courses. As a voice-to-text application, people with hearing disabilities could benefit from summarization to keep up with content in a more productive way.

An audio or text file will be taken as input to progress with this application, and data summarization processes will be implemented to get condensed text output.

## Introduction:

### Pipeline:



We will work on two phases in this:

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

Everyone is constantly exposed to voice apps and audio content, and often one needs to access such content quickly, just as we could do scanning through a written text, but of course, this is not possible with audio. So it would obviously be advantageous to summarize audio contents, be them dialogues, lessons, or speeches. That is why summarization is a reasonably significant concept in NLP.

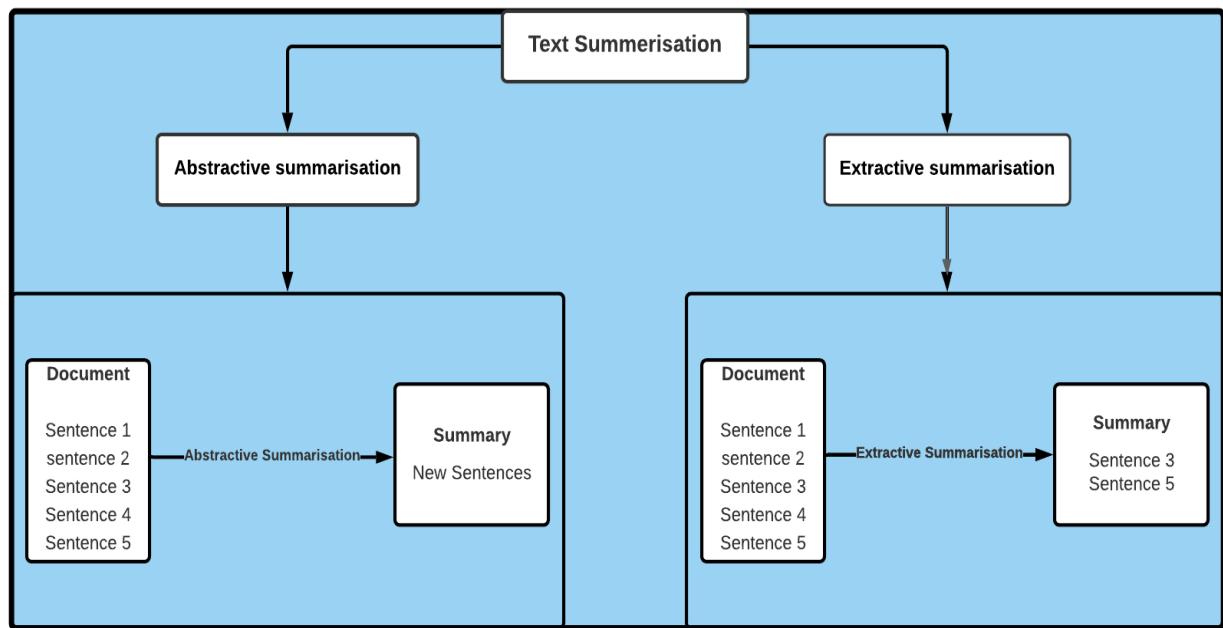
**“Perhaps the best test of a man’s intelligence is his capacity for making a summary”**

— Lytton Strachey

There are two common approaches that are widely used for text summarization:

**Extractive Summarization:** The model identifies the meaningful sentences from the input text and only outputs those as a summary.

**Abstractive Summarization:** The model produces an entirely different text shorter than the original. It generates new sentences in a new form, just like humans do. It is more challenging to create such a summary because it is semantic-based and training such a model is a complex task.



We will be performing both extractive and abstractive text summarization.

For an **extractive summary**, we will be using the NLTK toolkit like stopwords, vectorizing the sentence, and building a similarity matrix. Using PageRank we sorted the most important sentences and generated a summary.

For an **abstractive summary**, we will be comparing our custom fine-tuned model (Rouge-score) with multiple PEGASUS Models and show the best summary. (PEGASUS is a state-of-the-art model for abstractive text summarization developed by Google Research and published at the 2020 International Conference on Machine Learning).

## **Requirements**

<b>Req ID</b>	<b>Requirement Detail</b>
REQ1	The User should be able to upload an audio file for summarization with the given specifications.
REQ2	The User should be able to record audio for summarisation with given specifications.
REQ3	The User should be able to see the transcription of the imported or recorded file.
REQ4	The User should be able to upload the tex file for summarisation.
REQ5	The User should be able to download text/transcription of the audio file.
REQ6	The User should be able to see the different summaries of the audio transcription/text file.
REQ7	The User should be able to download the summary of the transcription/audio file or text file.

## **KDD For Audio to Text Summarization**

**1: Data Collection:** The model aims to create a coherent summary of the transcribed audio signals or text files.

So, in this case, the data will be collected by recording the audio signal from the UI meeting certain criteria or by uploading the text file for a summary. We can also upload the audio recordings which meet the required specification for the signal input.

We have used XSUM dataset to finetune Pegasus pretrained model, from Datasets library or we can download it from

<https://www.kaggle.com/aliisakroe/xsum-no-hallucinations>.

The dataset consists of 226,711 news articles accompanied with a one-sentence summary. The articles are collected from BBC articles (2010 to 2017) and cover a wide variety of domains (e.g., News, Politics, Sports, Weather, Business, Technology, Science, Health, Family, Education, Entertainment and Arts). The official random split contains 204,045 (90%), 11,332 (5%) and 11,334 (5) documents in training, validation and test sets, respectively.

**2: Preprocessing:** Since the input data is collected from abundant sources, it is crucial to preprocess data to improve the model's accuracy. Here we have used the Natural Language Toolkit for doing symbolic and statistical natural language processing of the input data to remove introductions and other stop words for the summarization model.

The sentences will be divided and then tokenized using the NLP toolkit in this stage.

### **3: Summarization Model:**

We have tokenized the input text for extractive summarization, applied NTLK stop words, and vectorized the data. Using cosine similarity in python, we built the similarity matrix and applied PageRank, selecting the critical sentences.

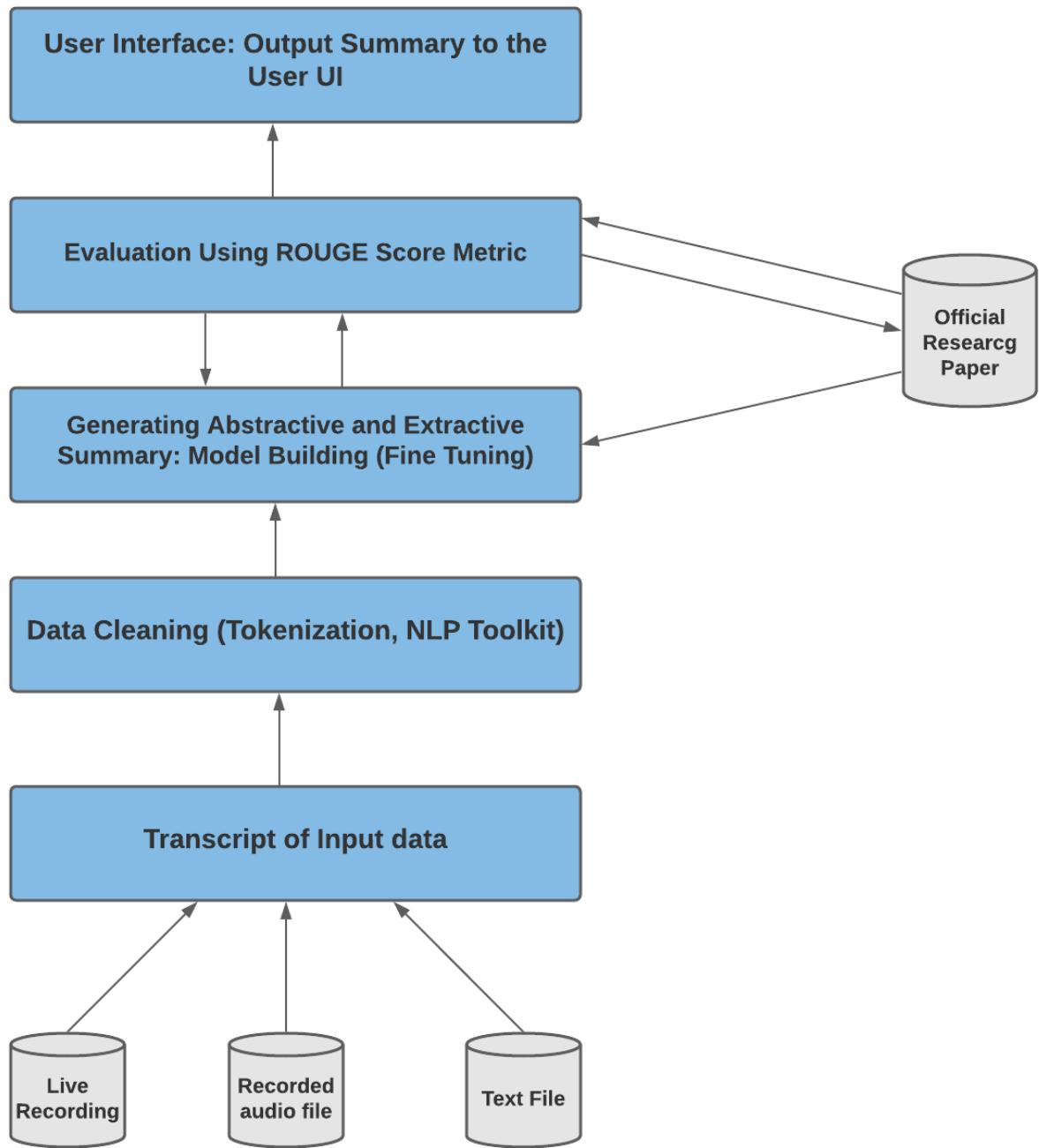
For abstractive summarization, the preprocessed transcript data is fed to PEAGUSUS model which will process the incoming transcript input and provide an abstractive summary. We have tuned the Pegasus model on XSUM dataset and deployed the model on AWS Sagemaker and Huggingface's official library.

The basic architecture of PEGASUS involves a standard Transformer encoder and decoder. It involves Gap sentence generation and the Masked Language model, in which a few words in the incoming input are masked and used as target generation text while the others are left as is.

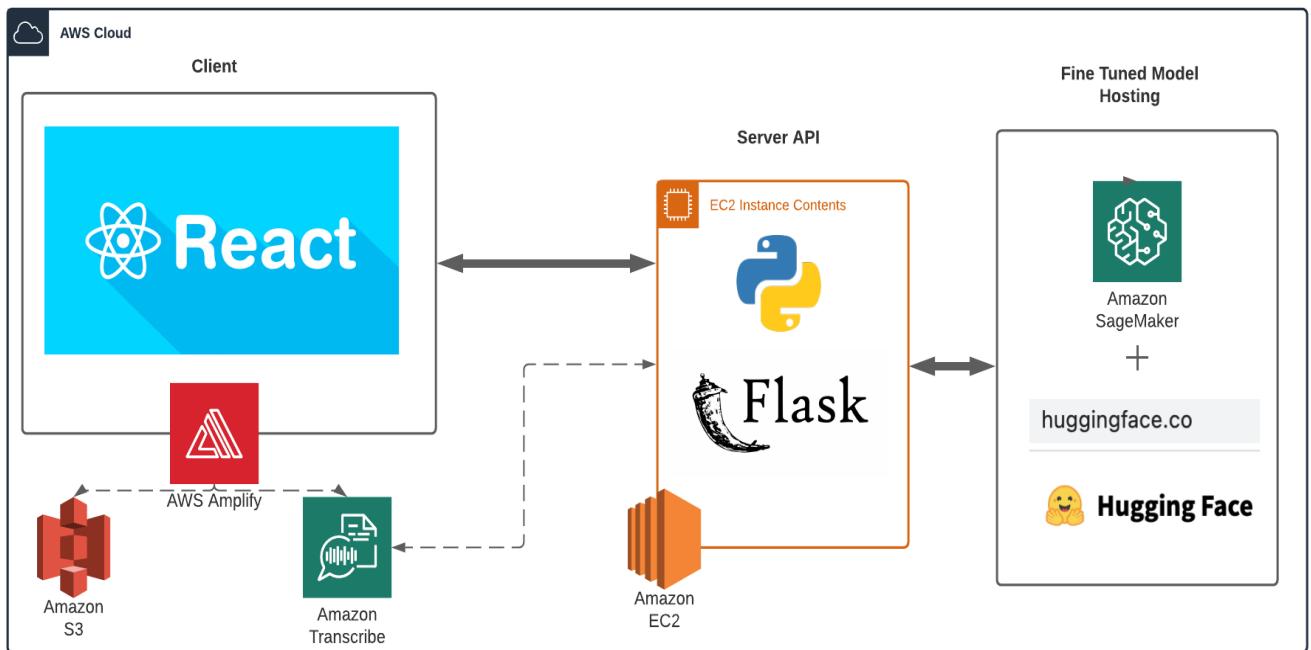
**4: Evaluation:** The output evaluation is done with a standard metric for abstractive summarization called ROUGE. It computes the similarity between 2 sentences by computing n-gram overlaps using a score from 0-100. ROUGE is a modification of BLEU that focuses on recall rather than precision. In our project, we used an Extractive summary as the reference summary, and the ROUGE score for all the other summaries is calculated against the extractive summary. The resulting window shows ROUGE scores for all the summaries used along with the name of the highest score summary model.

[source](#).

## High-Level Architecture Design and Data Flow Diagram:

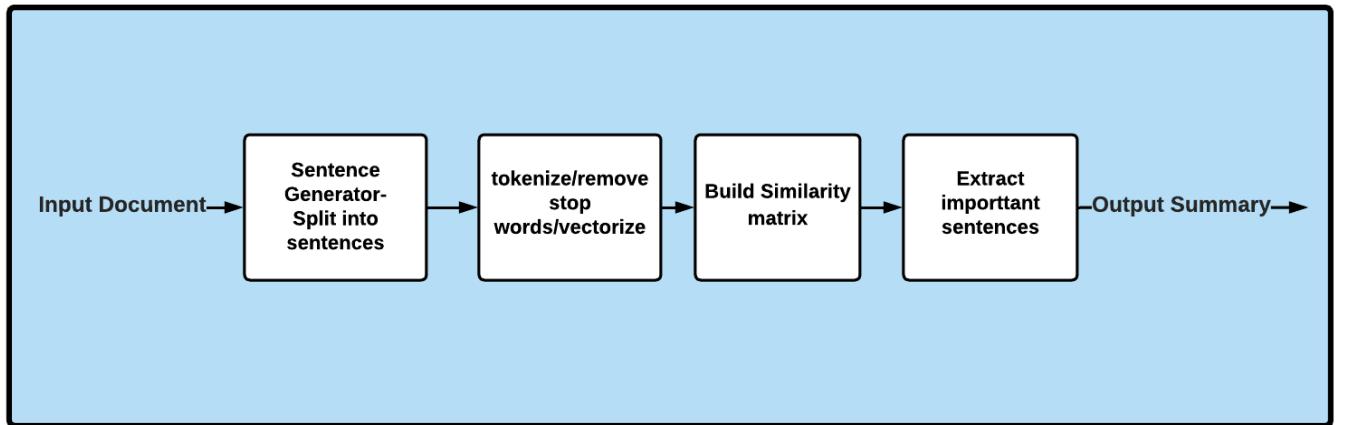


## Component Level Design:



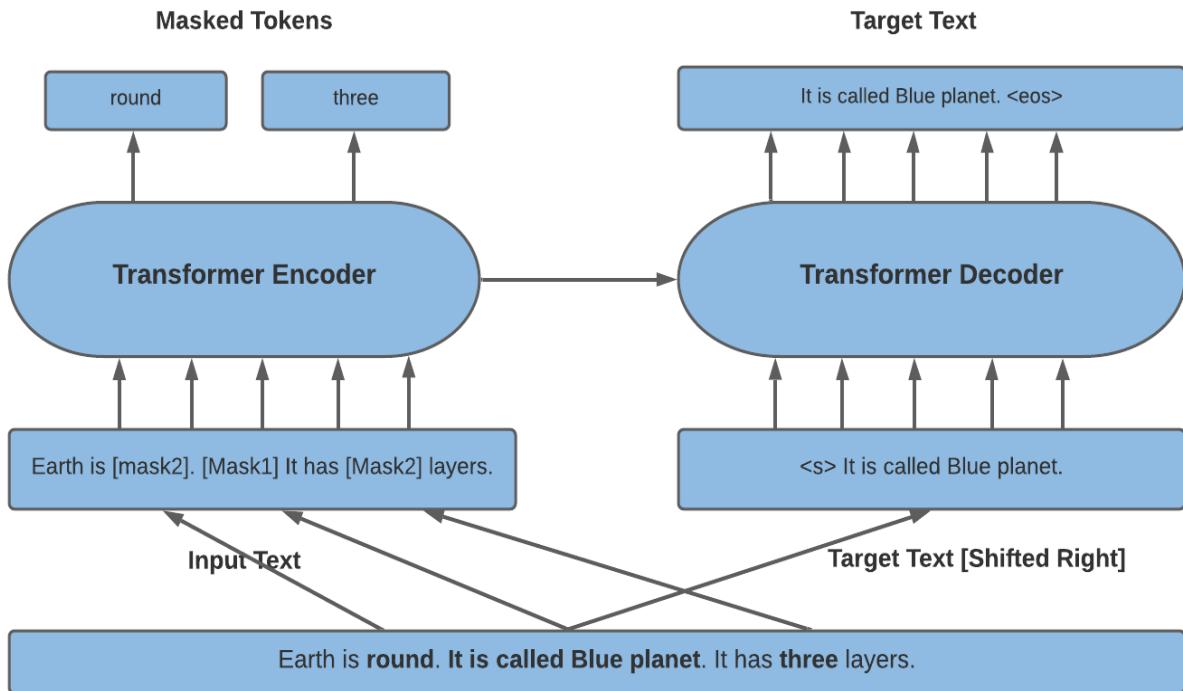
## Data Science Algorithm or Model Used

### Extractive Summarization:



### Abstractive Summarization:

Base Model -PEGASUS  
The architecture of the PEGASUS Model



The basic architecture of the PEGASUS model is based on a transformer encoder and decoder. In this model both GSG (Gap statement Generation) and MLM are applied simultaneously to the incoming data. The above figure shows the architecture of PEGASUS in which we have 3 sentences.

The first and last sentence remains in the input but some tokens in it are masked by [MASK2] (MLM). Whereas the middle sentence is completely masked with [MASK1] and it is used as target generation text (GSG). [source](#)

So basically in this model from the input statement, important sentences are masked/removed, and then an output sentence is generated from these statements. It is very similar to an extractive summary.

The text generated has to be finetuned later. Abstractive text summarization is said to be one of the most challenging tasks in NLP, this involves properly understanding long passages, compressing information and then language generation.

The most dominating paragon for training ML models to do summarization is Sequence-To-Sequence, in which the neural network masters to map the input sequences to output sequences.

Initially seq2seq models were developed using recurrent neural networks, on the other hand Transformer encoder-decoder models have lately become famous

and favorable, since they are very effective at modeling the dependencies which are present in the long sequences encountered in summarization.

In “PEGASUS: Pre-training with Extracted Gap-sentences for Abstractive Summarization” (to appear at the 2020 International Conference on Machine Learning), we designed a pre-training self-supervised objective (called gap-sentence generation) for Transformer encoder-decoder models to improve fine-tuning performance on abstractive summarization, achieving state-of-the-art results on 12 diverse summarization datasets.

On contrary to extracting the important sentences to generate a summary, in this technique novel sentences are generated by either rephrasing or using new words, then these sentences are then used to generate a summary thereby retaining the main crux of the document and being linguistically apt at the same time.

Researchers at Google very recently published PEGASUS (Pre-training with Extracted Gap-sentences for Abstractive SUmmarization sequence-to-sequence models) in the field of Abstractive text Summarization. This paper now stands as a State-of-the-Art model for Abstractive summarization when evaluated on multiple datasets.

At the starting of the paper, it points out all of the already existing methodologies which involve generalized pre-training self-supervised objectives that have been used by people on large aggregation which might not scale and perform efficiently when dealing with problems of text summarization.

A new pre-training objective specific to Text summarization is proposed in the paper which later could be fine-tuned and adapted to various complications with very little data. The model can provide good performance with as few as 1000 examples for the in-hand tasks.

A new Transformer based encoder-decoder pre-training objective is proposed in the paper which is called GSG(Gap Sentence Generation).

**GSG:** In this model certain sentences in the input document are masked and the remaining sentences are used to reproduce that sentence. The

respective position of each statement selected as a gap statement is replaced by a mask token [MASK1] to inform the model. GSR or Gap sentence ratio is referred to as the number of selected gap sentences to the total number of sentences in the document, which is similar to the mask rate in other works. To generate a more approximate summary, the sentences which are selected are more important/principle to the document. The paper also differentiated their main objective with other pre-trained models like MASS, UniLM, T5, and BART on the basis of selecting whole sentences rather than smaller continuous text chunks as done on the mentioned models.

The model overall revolves around 3 pre-training objectives: GSG, GSG+MLM (used both), and BERT's MLM (Masked Language Model).

Model along with GSG is used to train the transformer encoder model as the only pre-training objective. When MLM is used as the pre-training objective then the Transformer decoder shares all parameters with the encoder when fine-tuning.

Two large-scale corpora C4 and HugeNews are used for pre-training the model. The pre-trained model has been fine-tuned with 12 summarization datasets. These datasets belong to a wide variety of domains, for instance, tech, news, science, and vary in length too.

In this model, the author has trained two variants namely the base model and the large model. In which the base model has 12 transformer units for both encoder and decoder, on the other hand, the large model has 16 transformer units. The author has used greedy and beam search with length penalty as the sentence decoding strategy.

They could have tried other decoding strategies as well such as top-k, nucleus, and compare the performances. Neural Text Degeneration talks about the demerits of using Beam search as the choice for decoding.

They handle input sequences of length up to 1024 tokens. But they mention that since the average input length in BIGPATENT, arXiv, PubMed, and Multi-News are well beyond 1024 tokens, applying a two-stage approach (Generating Wikipedia by Summarizing Long Sequences) could

improve performance even more. The authors also talk about other hyper-parameters which you can read in the original paper. The below-mentioned table shows results on various datasets.

## **Server Side Components**

**Amazon Transcribe:** One of the initial tasks after loading the dataset was to convert the speech signal to text, so as to generate the summary in later steps. We have used Amazon Transcribe to deliver this part of the process. It is a very powerful AWS automatic speech recognition service that internally uses machine learning models to convert speech signals into text. It has certain features such as custom vocabularies in which we can provide custom words which we want the model to recognize, vocabulary filtering if we want to mask few words to not appear in the transcript.

**AWS Amplify:** It is a complete package which helps both the mobile and web developers to develop scalable and full stack applications which will run on AWS. This platform comes with number of tools and services that helps developer to integrate with applications, configure backend services, deploy any static website, manage backend contents, etc. It has a built-in CLI, ample code libraries, and already built-in components that can be used anytime.

**AWS Flask:** It is a web framework provided by AWS, which allows the developer to create a web application with a number of inbuilt libraries, tools, and technologies. It's a customizable python framework that helps to design a web framework for restful API development. In our project, we have used Flask to write API to run certain functionality requested from the frontend which includes running the model to generate summaries of different types.

## Client-Side Design

### Template Design



# Outcome

The screenshot shows a web browser window for the URL <https://dev.dp9zd0evuuig4.amplifyapp.com>. The page title is "Text Summarisation". The interface includes three input methods: "Audio File Input", "Text Input", and "Voice Input". The "Text Input" section contains a "Browse..." button, a message "No file selected.", and a placeholder "Upload your text file". The "Voice Input" section contains a "Start recording" button and instructions: "Press 'start recording' to begin your transcription. Press STOP recording once you finish speaking." Below these sections is a "Input Text" area with the placeholder "Waiting for input....". To the right of the input areas are two buttons: "Show Rogue Score Board" and "Print Score Card". On the far left, there are five vertical buttons, each labeled with a different model configuration:

- Generate Extractive Summary
- Generate 'google/pegasus-billsum' Abstractive Summary
- Generate 'google/pegasus-cnn\_dailymail' Abstractive Summary
- Generate 'google/pegasus-multi\_news' Abstractive Summary
- Generate 'arawat/pegasus-custom-xsum' Abstractive Summary

## Testing and Validations

1. The user has to either upload an audio, text file or record live audio from the below UI. The Input Text section shows waiting for input till the time the input file is loaded.

The screenshot shows a web browser window with the URL `localhost:3000`. The page title is **Text Summarisation**. The interface is divided into several sections:

- Audio File Input:** A file input field with the placeholder "No file selected." and a "Browse..." button. Below it is a text input field labeled "Upload your audio file".
- Text Input:** A file input field with the placeholder "No file selected." and a "Browse..." button. Below it is a text input field labeled "Upload your text file".
- Voice Input:** A button labeled "Stop recording". Below it is a text instruction: "Press 'start recording' to begin your transcription. Press STOP recording once you finish speaking."
- Input Text:** A large text area containing the placeholder "Waiting for input....".
- Action Buttons:** A row of buttons on the right side:
  - "Generate Extractive Summary"
  - "Generate 'google/pegasus-billsum' Abstractive Summary"
  - "Generate 'google/pegasus-cnn\_dailymail' Abstractive Summary"
  - "Generate 'google/pegasus-multi\_news' Abstractive Summary"
  - "Generate 'arawat/pegasus-custom-xsum' Abstractive Summary"
  - "Show Rogue Score Board"
  - "Print Score Card"

## Favorites

🕒 Recents

🕒 Applications

🕒 Desktop

🕒 Documents

🕒 Downloads

🕒 Pictures

## iCloud

🕒 iCloud Drive

## Locations

🌐 Network

## Tags

🔴 Red

🟠 Orange

🟡 Yellow

🟢 Green

🔵 Blue

&lt; &gt; ⌂ ⌂

articles

Search

Name

Size

Kind

msft.txt

2 KB Plain Text

bats.txt

4 KB Plain Text

pge.txt

281 bytes Plain Text

Cancel

Open

- Once the input file is loaded completely it will show the content of the file in the input text.

**Text Summarisation**

<b>Audio File Input</b>  Browse... speech.wav Upload your audio file	<b>Text Input</b>  Browse... No file selected. Upload your text file	<b>Voice Input</b>  Start recording Press 'start recording' to begin your transcription. Press STOP recording once you finish speaking.
-------------------------------------------------------------------------------	-------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------

**Input Text**

in an attempt to build an Iraqi work for Microsoft announced intelligent cloud hub which has been launched to him for the next generation of students with Iraqi skills envisioned as a three year collaborative program. Intelligent cloud hub will support around 100 institutions with I infrastructure, of course content and curriculum developer support development tools. End of students access to cloud. A nice 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 infrastructure and I hub for the selected campuses. The company will provide I development tools and Treasury services such as Microsoft cognitive services dot services and Azure Machine learning. According to Manish Prasad, Cash country general manager, ps Health and Education Microsoft India said with by being the defining technology of our time. It is transforming lives and industry and the jobs of tomorrow will require a different skill set. This will require more collaborations and training and working with I that's why it has become more critical than ever for educational institutions to integrate new cloud and I 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 new york, the company announced Microsoft professional program and I 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, and I and data science, with a series of online forces, which featured hands on labs and expert instructors as well. This program also included developer focused high school that provided a bunch of assets to help build I skills.

Generate Extractive Summary		Show Rogue Score Board
Generate 'google/pegasus-billsum' Abstractive Summary		Print Score Card
Generate 'google/pegasus-cnn_dailymail' Abstractive Summary		
Generate 'google/pegasus-multi_news' Abstractive Summary		

- We have used 4 types of summary generation techniques, the user can click on any of the buttons

In the buttons in the UI to generate the summary.

localhost:3000

50% ⌂ ⌄ ⌁

## Text Summarisation

<b>Audio File Input</b>	<b>Text Input</b>	<b>Voice Input</b>
<input type="button" value="Browse..."/> speech.wav Upload your audio file	<input type="button" value="Browse..."/> No file selected. Upload your text file	<input type="button" value="Start recording"/> Press 'start recording' to begin your transcription. Press STOP recording once you finish speaking.

**Input Text**

in an attempt to build an Iraqi work for Microsoft announced intelligent cloud hub which has been launched to him for the next generation of students with Iraqi skills envisioned as a three year collaborative program. Intelligent cloud hub will support around 100 institutions with infrastructure, of course content and curriculum developer support development tools. End of students access to cloud. A nice 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 infrastructure and hub for the selected campuses. The company will provide development tools and Treasury services such as Microsoft cognitive services dot services and Azure Machine learning. According to Manish Prasad, Cash country general manager, ps Health and Education Microsoft India said by being the defining technology of our time. It is transforming lives and industry and the jobs of tomorrow will require a different skill set. This will require more collaborations and training and working with that's why it has become more critical than ever for educational institutions to integrate new cloud and 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 new york, the company announced Microsoft professional program and I 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, and I and data science, with a series of online forces, which featured hands on labs and expert instructors as well. This program also included developer focused high school that provided a bunch of assets to help build I skills.

Generate Extractive Summary	in an attempt to build an Iraqi work for Microsoft announced intelligent cloud hub which has been launched to him for the next generation of students with Iraqi skills envisioned as a three year collaborative program. 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. The company will provide development tools and Treasury services such as Microsoft cognitive services dot services and Azure Machine learning. Earlier in april this new york, the company announced Microsoft professional program and I as a learning track open to the public. With the program will set up the core infrastructure and hub for the selected campuses	Show Rogue Score Board
Generate 'google/pegasus-billsum' Abstractive Summary		Print Score Card
Generate 'google/pegasus-cnn_dailymail' Abstractive Summary		
Generate 'google/pegasus-multi_news' Abstractive Summary		
Generate 'arawat/pegasus-custom-xsum' Abstractive Summary		

← → ⏪ localhost:3000 30% ⭐

## Text Summarisation

Audio File Input	Text Input	Voice Input
<input type="button" value="Browse..."/> No file selected. Upload your audio file	<input type="button" value="Browse..."/> msft.txt Upload your text file	<input type="button" value="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
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.	Print Score Card
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."	

4. Once the summaries are generated we can check the ROUGE Score using the button on the 'Show ROUGE Score Board' button. It shows the individual score along with the best score with model name. The complete summary page can be printed using the 'Print Score Card' button.

Text Summarisation		
Audio File Input	Text Input	Voice Input
<input type="button" value="Browse..."/> No file selected.  <input type="button" value="Upload your audio file"/>	<input type="button" value="Browse..."/> msft.txt  <input type="button" value="Upload your text file"/>	<input type="button" value="Start recording"/>  Press 'start recording' to begin your transcription. Press STOP recording once you finish speaking.
<b>Input Text</b> <p>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.</p>		
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	<input type="button" value="Show Rogue Score Board"/> <input type="button" value="Print Score Card"/> Extractive/Model Rouge =====> Value <b>google/pegasus-billsum =====&gt;</b> 0.3109243697478992  <b>google/pegasus-cnn_dailymail =====&gt;</b> 0.1818181818181818  <b>google/pegasus-multi_news =====&gt;</b> 0.4123076923076923  <b>arawat/pegasus-custom-xsum =====&gt;</b> 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.	<b>Best Score</b> Model Name: arawat/pegasus-custom-xsum Score: 0.5104602510460251
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.	

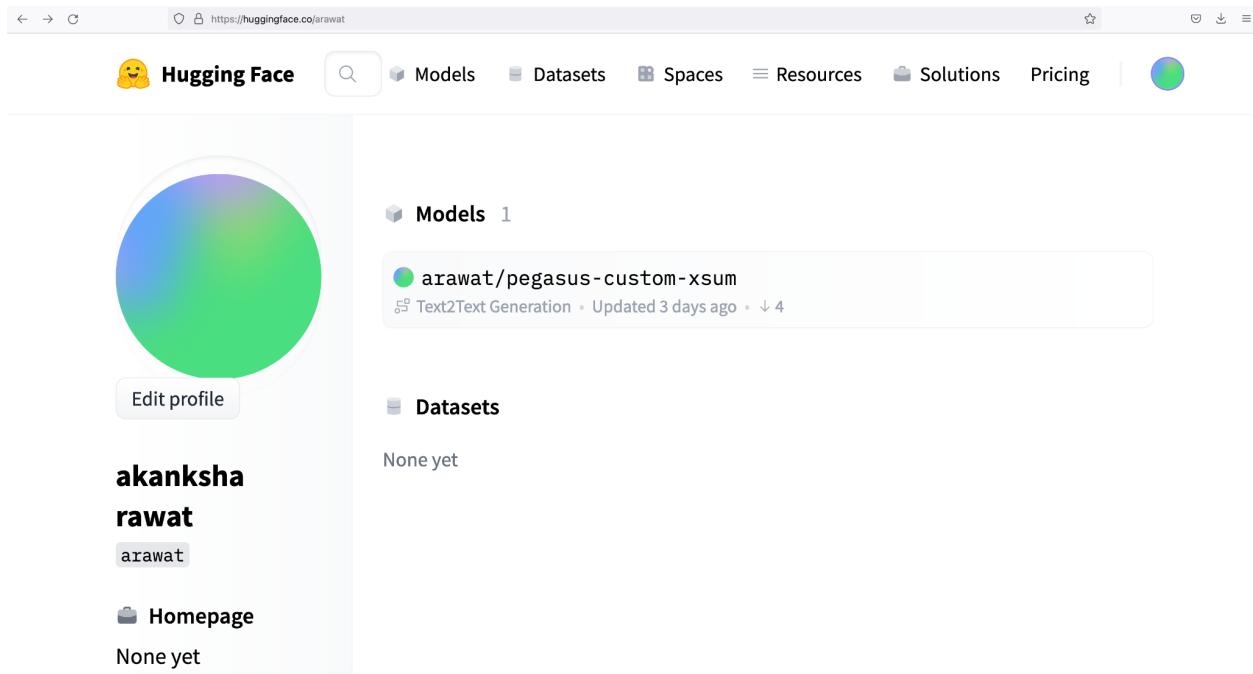
## Model Deployment

We have used Hugging face and Amazon sagemaker to finetune and deploy our custom model.

**Amazon Sagemaker:** It's an excellent service that facilitates data scientists to build, test, train and deploy complicated and high-quality machine learning models by providing a variety of capabilities purpose built-in for ML models. It also provides some standard models. Its solutions are fully customizable and easy to deploy with a single click.

Below are the details of the models deployed using huggerface and sagemaker.

- <https://github.com/akanksha0911/text-summarisation>
- <https://huggingface.co/arawat/pegasus-custom-xsum>
- [https://huggingface.co/arawat/pegasus-custom-xsum/blob/main/pytorch\\_model.bin](https://huggingface.co/arawat/pegasus-custom-xsum/blob/main/pytorch_model.bin)



Screenshot of the Hugging Face model card for `arawat/pegasus-custom-xsum`. The card includes sections for Model card, Files and versions, Settings, Train, Deploy, and Use in Transformers.

Screenshot of PyCharm showing the code for `DeploySageMaker.py`:

```
import requests
API_URL = "https://api-inference.huggingface.co/models/arawat/pegasus-custom-xsum"
headers = {"Authorization": "Bearer hf_aavWCpjNRS1qJfhlRJvDYNugutheEq"}
def query(payload):
    response = requests.post(API_URL, headers=headers, json=payload)
    return response.json()
output = query({
    "inputs": "In the distant past, many people thought bats had magical powers, but times have changed. Today, many people believe that bat wings are much like those in arms and the human hand, with a thumb and four fingers. In bats, the bones of the arms and the four fingers of the hands are very long. Since most bats stay in darkness during the day and do their feeding at night, they do not use their vision to maneuver in the dark but use a process called echolocation."
})
print(output)
```

The Run tab shows the output of the script:

```
/Users/akanksha/PycharmProjects/text-summarisation-final/venv/bin/python /Users/akanksha/PycharmProjects/text-summarisation-final/DeploySageMaker.py
[{'generated_text': 'Today, many people believe that bats are rodents, that they cannot see, and that they are more likely than other animals to carry rabies. The bones in bat wings are much like those in arms and the human hand, with a thumb and four fingers. In bats, the bones of the arms and the four fingers of the hands are very long. Since most bats stay in darkness during the day and do their feeding at night, they do not use their vision to maneuver in the dark but use a process called echolocation.'}]
```

A message at the bottom right of the PyCharm interface says: "Looks like you're using NumPy. Would you like to turn scientific mode on? Use scientific mode Keep current layout..."

arawat | pegasus-custom-xsum

Text2Text Generation PyTorch Transformers pegasus seq2seq generated\_from\_trainer model-index AutoNLP Compatible

Model card Files and versions Settings Edit model card Train Deploy Use in Transformers

results

This model is a fine-tuned version of google/pegasus-large on an unknown dataset.

Downloads last month 4

Hosted inference API

Text2Text Generation In an attempt to build an AI-ready workforce, Microsoft Compute Computation time on cpu: 34.5036 s

Model description More information needed

Intended uses & limitations More information needed

Training and evaluation data More information needed

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

arawat | pegasus-custom-xsum

Text2Text Generation PyTorch Transformers pegasus seq2seq generated\_from\_trainer model-index AutoNLP Compatible

Model card How to serve this model with the Accelerated Inference API Select Token Copy to clipboard Use in Transformers

results Try the Inference API for free, and get an organization plan to use it in your apps.

```
import requests
API_URL = "https://api-inference.huggingface.co/models/arawat/pegasus-custom-xsum"
headers = {"Authorization": "Bearer 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"}
```

Model description More information needed

Intended uses & limitations More information needed

Training and evaluation data More information needed

Inference API Documentation Pin this model for instant availability How to get started

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

## Active Learning

Since we have used multiple models to generate summaries, we got an opportunity to develop an extractive summary from scratch. We have understood and fine-tuned one SOTA model for an abstractive summary. We got exposure to deploy a custom model and fine-tune the model. We learned the functioning of all these models.

Text summarization has always been the most challenging task in NLP and has been ongoing for a long time now. Several models are providing less or more accurate summaries. Since not just the text but the context of the input has to be well understood by the model, it makes it much more difficult to generate a summary. During the development course of the project, we managed to learn the deployment, training, and fine-tuning of the custom model. While implementing and evaluating these models, we could understand how the result generated could be used as feedback to retrain the model to provide accurate results next time. We learned how to evaluate the performance of the models based on the ROUGE Score and compare them to discover the most accurate one. Especially with the Pegasus model, we got to understand the overall functioning of the model, the use and function of the transformers with encoders and decoders.

The key learning was the difference in extractive and abstractive summarization. How abstractive summarization is a challenging task. For abstractive summarization, the pre-training of the model is closer to the self-supervised objective to the final downstream task. It is challenging to train abstractive models on huge, billions of datasets to generate a human-like summary, so it is better if we use a pre-trained model and fine-tune that for our use case.

1. [Presentation link.](#)

2. GitHub Repo link:

<https://github.com/akanksha0911/text-summarisation-webapp>

<https://github.com/akanksha0911/text-summarisation>

3. References/Source:

<https://medium.com/analytics-vidhya/pegasus-pre-training-with-extracted-gap-sentences-for-abs-tractive-summarization-acb238aa1096>

<https://arxiv.org/pdf/1912.08777.pdf>

4. Model Deployment:

<https://github.com/akanksha0911/text-summarisation>

[https://huggingface.co/arawat/pegasus-custom-xsum/blob/main/pytorch\\_model.bin](https://huggingface.co/arawat/pegasus-custom-xsum/blob/main/pytorch_model.bin)

<https://huggingface.co/arawat/pegasus-custom-xsum>