MINUTES OF MEETING

Team Members

Seif Eldin Haybat

Basma Adawy

Mayar Gamal Ahmed

Ahmed El-Saidy

AGENDA

- Define Project
- Data
- Methodology
- Meeting Minutes Modules
  - Speech to Text
  - Meeting Summarization
  - Meeting Action Items and Decisions
- Deployment
- Future Work
- Conclusion

Aren't you tired of taking notes at every meeting you take?

WHAT ARE MEETING MINUTES?

Meeting Minutes, or informally, notes, are the record of a meeting in written form.

They typically describe the events of the meeting and may include a summary, a list of attendees, a listing of issues and decisions made, and an action list with due dates.

The minutes of a meeting are usually taken by a designated

BUSINESS CHALLENGES

01

Too much information is given by participants, Should he note them all down and hope for the best? Should he religiously expunge all discussions irrelevant to the main topic from the minutes**?.**

02

Participating in the call versus minute-taking, should he focus on the minute-taking or put the pen down and the figurative pointer up**?.**

PROJECT GOALS

- **Reduce the effort of manual Minutes of Meeting preparation to the maximum extent possible.**

- **Release a structured document of the notes for the participants of the meeting so that the attendees can completely focus on the discussion.**

DATASET USED

AMI Dataset

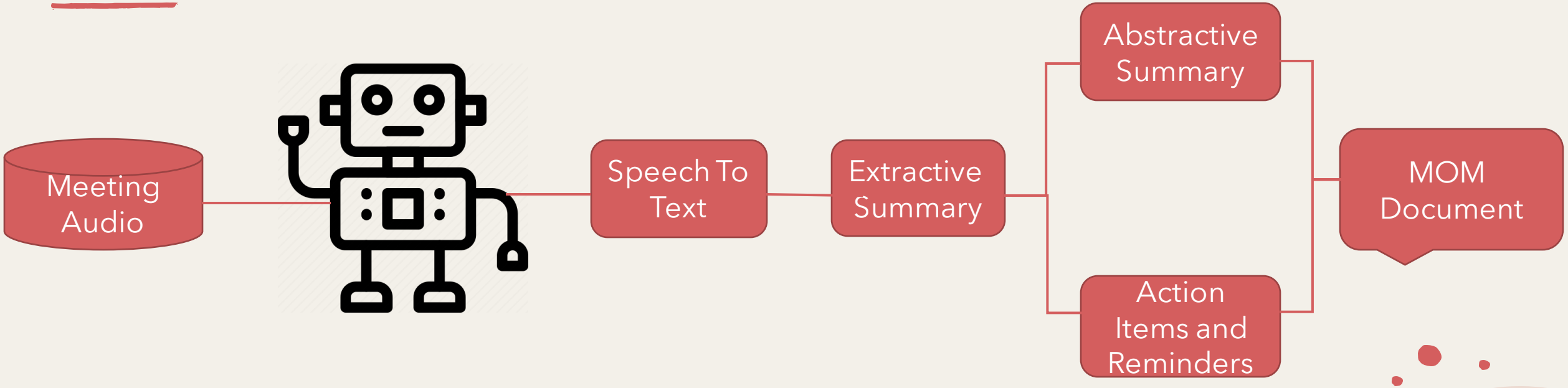
- The **AMI** Meeting Corpus consists of 100 hours of meeting recordings.
- The dataset supports:

1. Transcripts.
2. Extractive and Abstractive Summaries.
3. Action Items

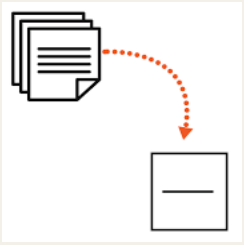
TEXT PREPROCESSING

- Convert text to lowercase
- Handling comma's ('s)
- Handling disfluencies
- Mapping contraction (can't: cannot)
- Remove punctuations and special characters
- Handling stopword
- Text stemming

MEETING MINUTES OVERVIEW



TECHNOLOGIES



Speech to Text	Extractive Summarization	Abstractive Summarization	Action Item Detection
Google Speech Recognition	Gensim TextRank	T5 Transformer	Rule Based Model

TEXT

SPEECH RECOGNITION

Python provides an API called **Speech Recognition** to allow us to convert audio into text for further processing.

The SpeechRecognition library : acts as a wrapper for several popular speech APIs and is thus extremely flexible. One of these– the **Google Web Speech API**–supports a default API key that is hard-coded into the Speech Recognition library. That means you can get off your feet without having to sign up for a service .

https://realpython.com/python-speech-recognition/

PROCESSING LARGE AUDIO FILES

When the input is a **long audio** file, the **accuracy** of speech recognition **decreases**. Moreover, Google speech recognition API cannot recognize long audio files with good accuracy.

Content Limit	Audio Length
Synchronous Requests	~1 Minute
Asynchronous Requests	~480 Minutes*
Streaming Requests	~1 Minutes

https://wicg.github.io/speech-api/
https://www.youtube.com/watch?v=LTOu9_IWMyQ&ab_channel=LINEDevelopers

SPLITTING THE AUDIO

split the audio file into chunks of **constant size**

split the audio file based on **silence** using pydub

https://ngbala6.medium.com/audio-processing-and-remove-silence-using-python-a7fe1552007a

HOW TO EVALUATE SPEECH TO TEXT?

Word Error Rate is a measure of how accurate an Automatic Speech Recognition (ASR) system performs.

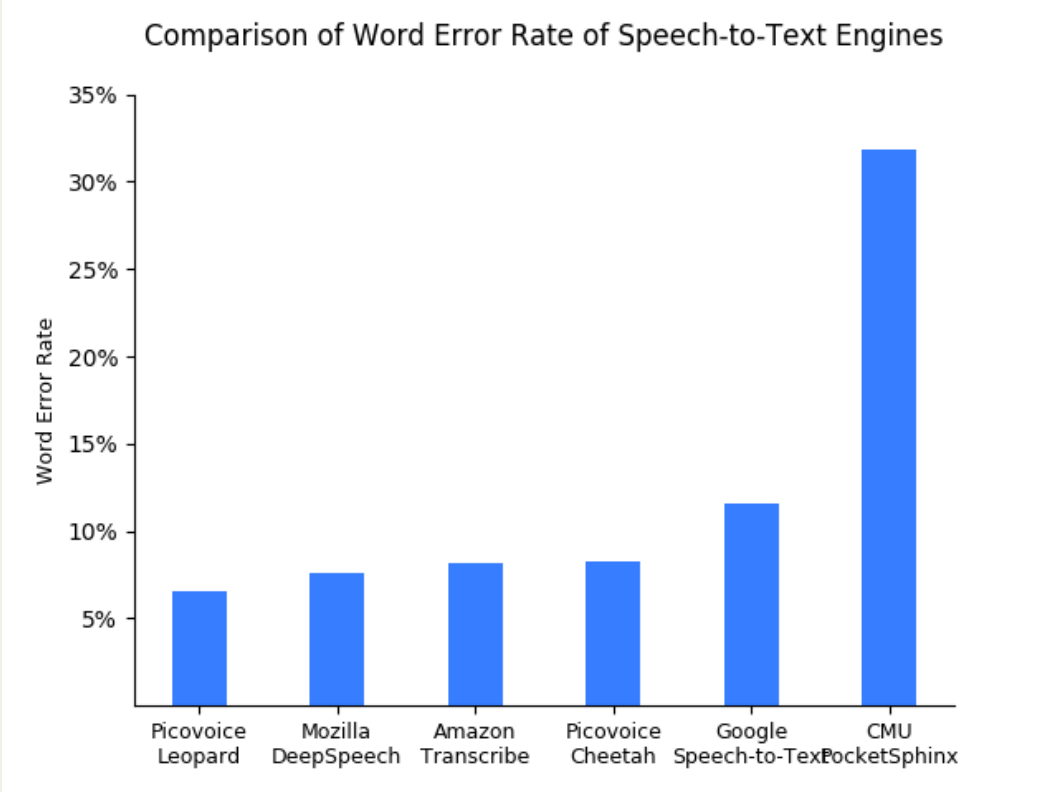
$$WER = \frac{S + D + I}{N}$$

of **Substitutions (S)**, **Deletions (D)**, and **Insertions (N)**, divided by the **Number of Words (N)**.

when computing a WER, through things like:

- Lowercasing all text
- Removing all punctuation
- Changing all numbers to their written form ("7" -> "seven")

https://github.com/Picovoice/speech-to-text-benchmark#data



LibriSpeech dataset
is used for benchmarking

Word Error Rate: 29.2%

Place Manual Transcription text in here

What about electric planes? Yes. We have a lot on our plate here. But electric planes, yes, I have been dying to do that for a decade, honestly. But we got to – we got quite a few fish to fry here. So, maybe one day the electric plane, battery energy density is improving every year. So, that's an important metric to get the sell energy density to around 450, 500 watt hours per kilogram and have a pack efficiency of around 400-watt hours per kilogram, that's when electric planes start to get interesting. So, it would be a fun problem to work on at some point. But we have a lot to do over the next few years. So, we are going to focus on these things, get them right and maybe one day to do that.

Alright. Thanks everyone for tuning in. Thanks for being here.

Place Automatic Transcription text in here

pack will be down low and. Maybe we can do some things with the suspension it just make it really hard to roll this thing. So it'll be in a tbs rollers when bad things happen. Cybertruck. Okay so i'll take one last question. Electric planes. Oh yeah yeah we have a lot on my plate here electric planes. But we we got. Read quite a few fish to fry here. So. Everyone day. B18 electric lane. A battery energy density is improving every year so that's an important. Ford metric get the. So anderson sc to. Around for 5500 watts to kg. And habitats have a pack efficiency around. 401000 kg that's when electricland saw to get to get interesting. Informant working at some point. But i. Sweet we got alexa over the next few years so we got to focus on. Love you thanksgiving right and then maybe one day do that. Thanks for opinion.

Word Error Rate: 50.0%

Place Manual Transcription text in here

hello dear

Place Automatic Transcription text in here

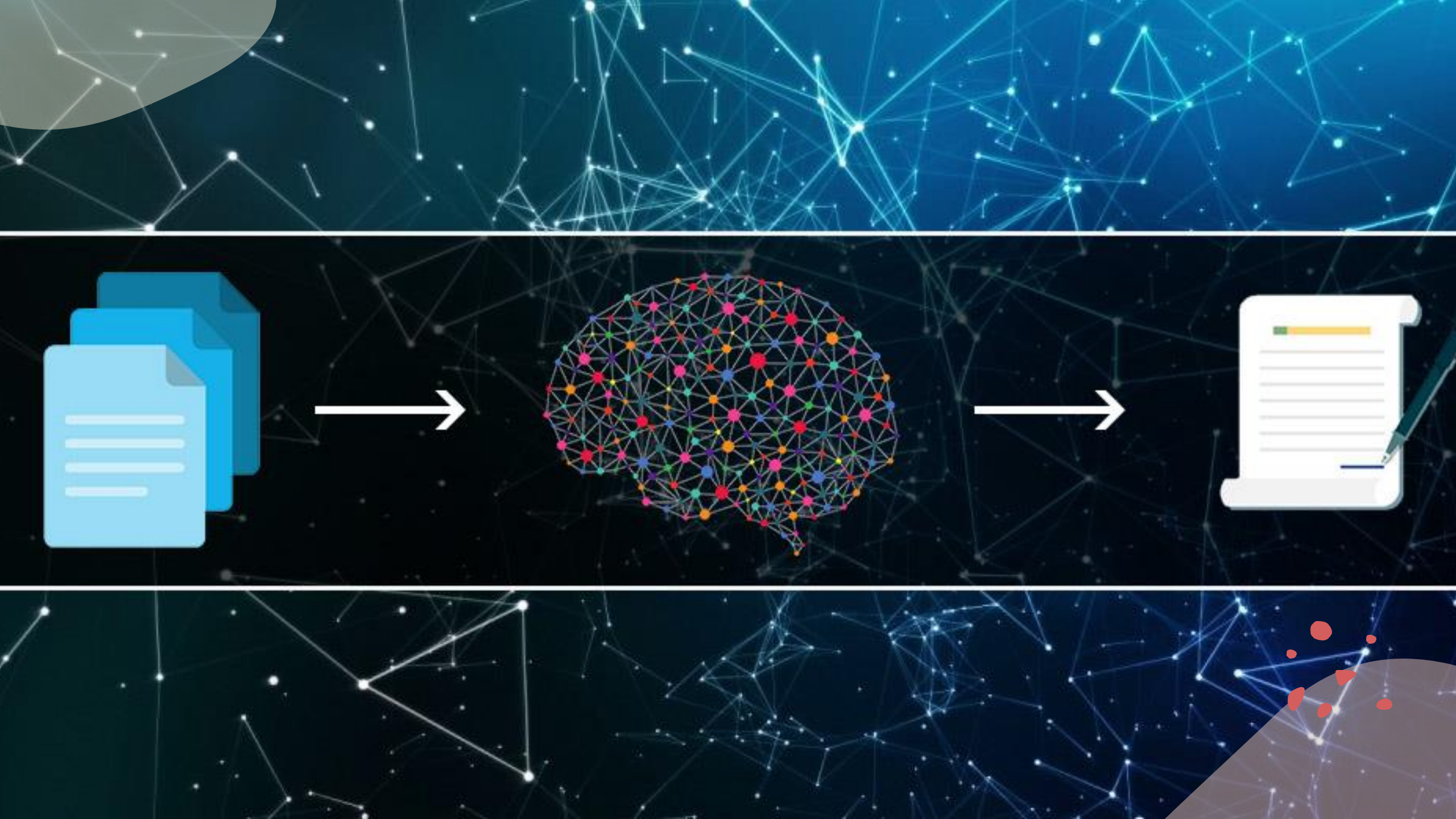
hello bear

https://www.amberscript.com/en/wer-tool

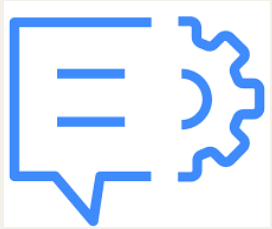
"You can have two systems with similar accuracy rates that produce wildly differently transcripts in terms of understandability. You can have two different systems that are similar in terms of accuracy but maybe one handles particular vocabulary that's germane to your application better than the other. There's more than just accuracy at the heart of it."

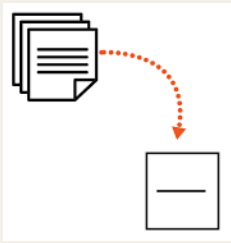
Klint Kanopka
Stanford Ph.D. Researcher

https://en.wikipedia.org/wiki/Speech_recognition

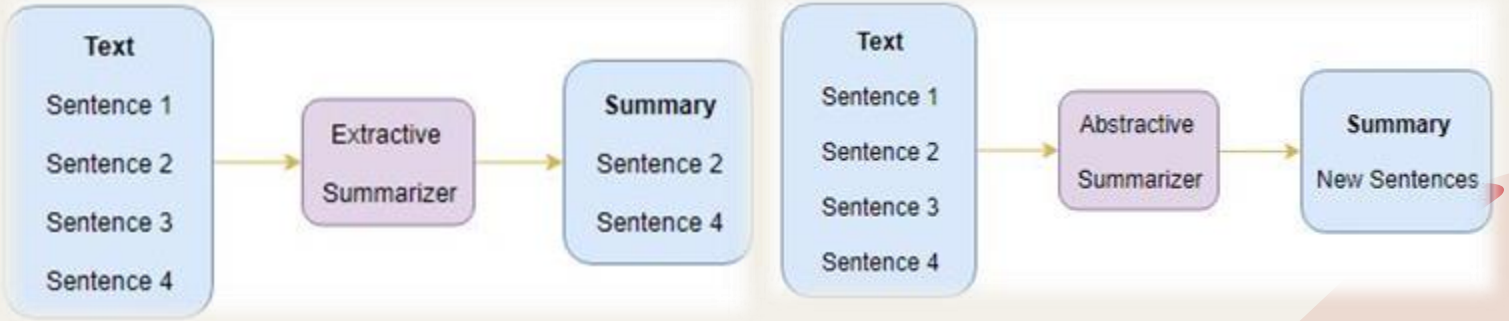


SUMMARIZATION MODULE





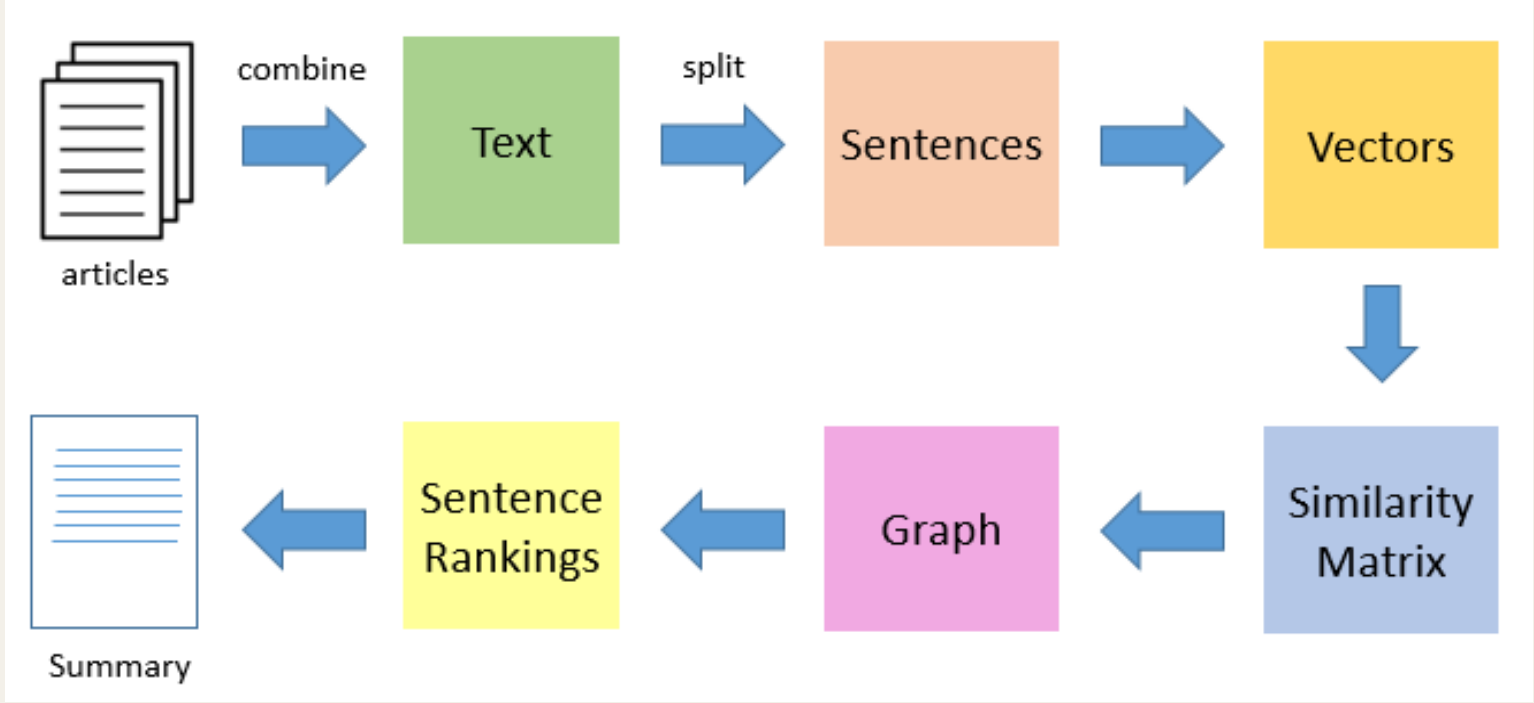
AMI Dataset | **Text Pre-Processing** | **Extractive Summarization** | **Abstractive Summarization** | **Evaliuation**



Gensim TextRank Extractive Summarization

Gensim TextRank

- Gensim is a free Python library based on the popular TextRank algorithm.
- TextRank is graph based model.

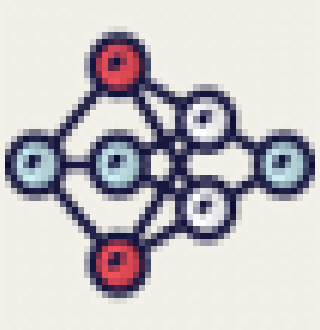


Evaluation

- **ROUGE**-N measures the number of matching 'n-grams' between our model-generated **text** and a 'reference'.

	rouge-1	rouge-2	rouge-l
r	0.690181	0.498344	0.677453
p	0.783857	0.647229	0.768870
f	0.726767	0.551317	0.713124

T5 Abstractive Summarization



Model	Challenges	Solution	Results
T5	Meetings are very long and all models have limited max length.	Double Summarization	ROUGE 18% **state-of-the-art** 24%

Abstractive Summarization

AMI abstractive summary

```
'The project manager introduced the upcoming project to the team members and then the team members participated in an exercise in which they drew their favorite animal and discussed what they liked about the animal. The project manager talked about the project finances and selling prices. The team then discussed various features to consider in making the remote.'
```

BART abstractive summary

```
'My favourite animal is like A beagle . , right , well basically high priority for any animal for me is that they be willing to take a lot of physical affection from their family. According to the brief we are gonna be selling this remote control for twenty five Euro , and we are aiming to make fifty million Euro . And we do not want it to cost any more than twelve fifty'
```

T5 abstractive summary

```
'Laura is the project manager and she is designing a new remote control. The project is aiming to make fifty million Euro. The remote control will be sold for twenty five Euros.'
```

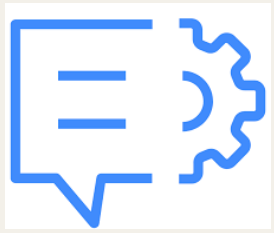
MEETING

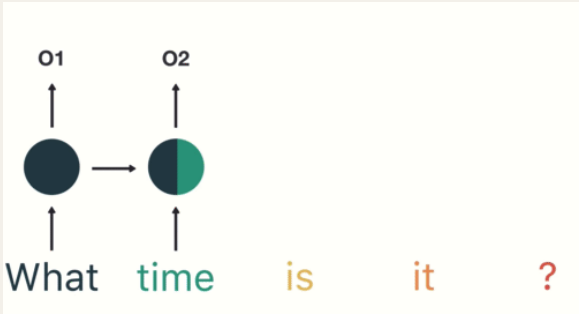
ACTION ITEMS

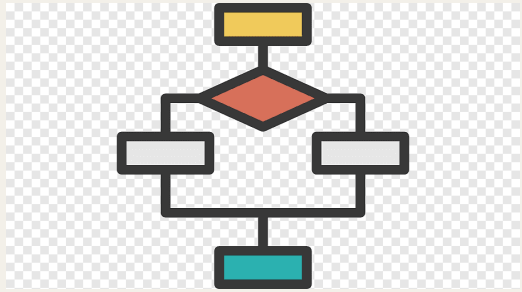
What is ACTION ITEMs ?

- An action item is a task that is created from a meeting with other stakeholders to move a project towards its goal.
- This action item is typically born from meeting notes and meeting tasks that arise over the course of the discussion.
- Writing action items doesn't have a clear standard
- We will focus on two action items : **Decisions & Reminders**

ACTION ITEM DETECTION – Decisions
Rule-Based Model







Transcript	DistilBERT Summarization	Tokenization & POS tagging	Rule-Based Model	Evaliuation
		Split summary into Sentences. Then tag each word in different category	Classify each sentence based on POS tagging and keywords	

• ACTION ITEM DETECTION – Decisions DistilBERT Summary

- BERT was introduced by Google in 2018 and since then it has shown state-of-the-art results in different language understanding tasks.
- DistilBERT is a small, fast, cheap and light Transformer model trained by distilling BERT base.
- *knowledge distillation during the pretraining phase and show that it is possible to reduce the size of a BERT model by 40%, while retaining 97% of its language understanding capabilities and being 60% faster.*

Transcript	DistilBERT Summary
We could do a better job at vetting resumes	Tesla should do a better job of vetting resumes.

ACTION ITEM DETECTION – Decisions Rule-Based Model

- DistilBERT Summary

> Jeff wants to train a Transformers model on Amazon SageMaker. He can use the new Hugging Face Deep Learning Container. The documentation is available on HuggingFace.co and on the blog, Jeff can find it here. Jeff will train the Transformers model with the new container.

ACTION ITEM DETECTION – Decisions Rule-Based Model

- Tokenize sentences and Tagging Words using NLTK
- A Part-Of-Speech Tagger (POS Tagger) is **a piece of software that reads text in some language and assigns parts of speech to each word**

```
[('Tesla', 'NNP'),
 ('should', 'MD'),
 ('do', 'VB'),
 ('a', 'DT'),
 ('better', 'JJR'),
 ('job', 'NN'),
 ('of', 'IN'),
 ('vetting', 'VBG'),
 ('resumes.\u200b', 'NN')]
```

JJR	adjective, comparative (larger)
JJS	adjective, superlative (largest)
LS	list market
MD	modal (could, will)
NN	noun, singular (cat, tree)
NNS	noun plural (desks)
NNP	proper noun, singular (sarah)

RB	adverb (occasionally, swiftly)
RBR	adverb, comparative (greater)
RBS	adverb, superlative (biggest)
RP	particle (about)
TO	infinite marker (to)
UH	interjection (goodbye)
VB	verb (ask)

ACTION ITEM DETECTION – Decisions Rule-Based Model

- based on tagging and keywords

```
def get_chunks(tagged_sent):
    chunkgram = r"""Modal-Auxiliary: {<NN><MD><VB>}   #Will May Might
                    Modal-Auxiliary: {<NN><MD><RB><VB>} # Will not
                    VB-Phrase: {<NN><VBZ><TO>}  #Decides to
                    VB-Phrase: {<NN><VBD><TO>}  #Decided to
                    ING: {<VBG><TO>}   #Going to
                    """
```

- Results:

	rouge-1	rouge-2	rouge-l
r	0.220291	0.042325	0.200958
p	0.333867	0.086937	0.304922
f	0.233489	0.046779	0.212394

ACTION ITEM DETECTION – Decisions T5 Fine-Tuning

- T5 summarization Model
- AMI Dataset has human written decisions
- Fine-tuning t5 on AMI Dataset decisions
- Results:

	rouge-1	rouge-2	rouge-l
r	0.152222	0.015790	0.140752
p	0.229804	0.024021	0.213618
f	0.176478	0.017875	0.163506

ACTION ITEM DETECTION - Decisions

Results

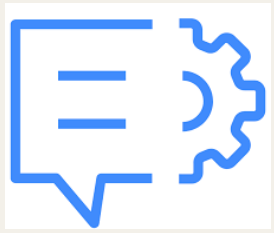
1. T5 Fine-Tuning

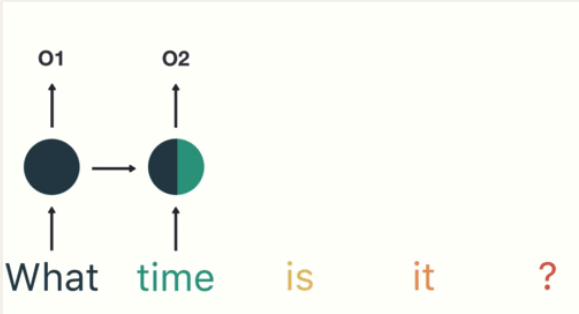
```
'We have The The The Tesla. The company has Tesla. The company has Tesla. The company has a vision of 20 million electric vehi
cles. The company has The formal part will cover the nine items that stockholders will vote on. The technical part will cover
the nine items that stockholders will vote on. The formal part will cover the nine items that stockholders will vote on. The f
ormal part will cover the nine items that stockholders will vote on. The company has committed to electric. Tesla.'
```

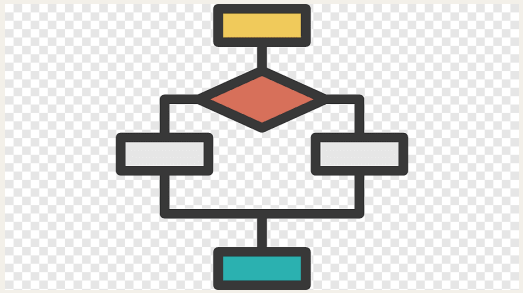
2. Rule-Based Model

```
['Tesla cannot rest on its laurels assuming its first-mover advantage will last.',
 'Model y will be the best selling vehicle of any kind globally next year.',
 "They are going to write an impact report on the company's impact report.",
 "Tesla's long-term competitive advantage will be.",
 'Tesla should do a better job of vetting resumes.The 1042-s are going to start rolling out tomorrow at midnight California time.',
 'Electric power will probably double for utilities.']
```

ACTION ITEM DETECTION – Reminders Rule-Based Model







Transcript	DistilBERT Summarization	Tokenization & POS tagging	Rule-Based Model
		Split summary into Sentences. Then tag each word in different category	Classify each sentence based on POS tagging and keywords

Case Study

We chose TESLA Stockholders annual meeting

Recorded in October 2021

GitHub

CLICK HERE

FUTURE WORK

- Minutes of Meeting on Arabic Meetings.
- Email server for sending reminders and decisions to attendees.

Now imagine if you have this technology in your hand?

THANK YOU!