Workshop1

Author: Paramate Phuengtrakul
Student ID: 640631124

Project(app) name: MinuteMade — Your meeting minutes, made better and faster.

Business Problem

- Teams often struggle to review long meeting records or notes.
- Manual note-taking is time-consuming and often leads to mistakes or missed detail

Remark: It's the same context problem (one problem), it's just separated into sub-task problems

Business Objective

Develop a software product that provides automated meeting summaries by integrating with popular meeting platforms. This helps teams save time, improve productivity, and ensure clear communication.

Activities, outputs, outcomes, and indicators identification

Activities	Outputs
Apply Natural Language Processing (NLP) techniques to generate concise, meaningful summaries.	Summarized meeting notes.
Integrate with meeting platforms (e.g., Microsoft Teams, Zoom, Google Meet) via APIs. Or implement to Web-app	Meeting platform APIs: Automatic retrieval of meeting transcripts and metadata. Web-app: Users are able to drag and drop audio file and retrieve summarized text easily

Outcomes	Indicators
Reduced time spent reviewing meeting content.	Reduction in time spent (measured via Full-Time Equivalent [FTE]).
Better alignment on tasks and decisions within teams.	Number of decisions/action items successfully executed.
Increased user satisfaction with meeting-related workflows.	User satisfaction score (e.g., Net Promoter Score or surveys).
Enhanced team collaboration with shared meeting insights.	Number of users actively sharing or using summaries.

Stakeholders

Stakeholder group	Who	Responsibility
Business Stakeholders	Managers Executive Department Head	Define goals, approve budgets.
End Users	Employees/Department (who interests)	Provide feedback, validate functionality.
Development Team	Frontend Developer Backend Developer Data Engineer Data Scientist	Build and maintain the software.
Vendor (AWS Partner or other cloud providers)	Vendors	Manage AWS services, provide support, ensure scalability.
Legal team	Legal Department	Ensure compliance with laws and regulations. (e.g., PDPA)

Data Source

MinuteMade, involved 2 models, namely, Speech-to-Text and Text Summarization,

- **Speech-to-Text Dataset**
  - Meeting platforms: APIs from Zoom, Microsoft Teams, Google Meet, etc.
  - Audio transcript by third-party APIs (e.g., Google, OpenAI), that's generated from input video meeting or seminar context from Youtube or other video streaming platforms.
  - Raw data features
    - audio file (e.g., .wav, .flac, .mp3, .m4a)
    - timestamped transcripts (texts) (e.g., .txt, .json, .csv, .xlsx)

- **Text Summarization Dataset**
  - Internal document company
  - Provided dataset in AI-community (e.g. Kaggle, HuggingFace, paperswithcode)
    - e.g., PUBMED
      - 19717 scientific publications from PubMed database
  - Raw data features
    - Text file (e.g., .json, .txt, .csv)
      - all contexts, and its summarization

Data Mining Technique

- **Text processing**
  - clean
    - normalize (e.g., converting to lowercase)
    - Remove filler words (e.g., "uh," "um").
    - Eliminate special characters, URLs, or unnecessary formatting.
  - tokenize
    - pre-trained tokenizers (e.g., T5, GPT-2)

- **Signal processing**
  - Extract feature (e.g., Fourier Transform, MFCC: *Mel-frequency cepstrum coefficients)*
  - Normalize feature vectors

- **Speech-to-Text models**
  - Deep Learning technique
    - The combination of CNN and RNN architecture
      - CRDNN with CTC/Attention and RNNLM model
  - Third-party API
    - Whisper by OpenAI

- **Text Summarization models**
  - Deep Learning technique
    - Transformer architectures
      - Fine-tune on T5 (Text-to-Text Transfer Transformer)
      - Fine-tune on GPT-2
  - Third-party API
    - GPT-4o by OpenAI

Data Product (Solution)

Web Application
Name: MinuteMade — Your meeting *minutes*, *made* better and faster.

- **User Log-in**
  - Authenticate via Microsoft Teams API to fetch meeting metadata.
- **Fetch or Upload**
  - User selects a meeting from their Teams history or uploads an audio file.
- **Speech-to-Text**
  - Transcribe the audio using Whisper or CRDNN.
- **Text Summarization**
  - Generate a concise summary using T5 or GPT-2 or GPT-4o
- **Display Results**
  - Show the summary on the dashboard with options to download, share, or copy.

System Architecture

Response transcript text

Microsoft Team

Store activities log

RDS from AWS

Call MS team API

Store historical results

request

response

User

Web application

Audio file

Text-to-Speech model API

text

Text Summarization model API

summarized text

Fetch model

Fetch model

Frontend image

CRDNN with CTC and RNNLM

Whisper by OpenAI

T5 GPT-2

GPT-4o by OpenAI

push image to ECR

Model image

pull image from ECR

push image to ECR

ECS from AWS

ECR from AWS