

Machine Learning Group20 Fifth Meeting Minutes					
Place	Google Meet ( <a href="https://meet.google.com/ztz-gsbv-ojw">https://meet.google.com/ztz-gsbv-ojw</a> )				
Time	2023/11/20 (Mon.) 20:00 ~ 21:30				
Minutes Writer	楊書蓉				
Minutes Refresher	吳姿儀				
Agenda					
Time spot	Discussion topic			Presentation	
20:00 ~ 20:15	Text process in ASR			林宸宇	
20:15 ~ 20:30	Text process in NPL			張睿云	
20:30 ~ 20:45	Line-Bot building & APIs			楊書蓉	
20:45 ~ 21:00	Data process in MFCCs			吳姿儀	
21:00 ~ 21:15	Data splitting issues			黃子軍	
In attendance					
Teammate	林宸宇	張睿云	吳姿儀	楊書蓉	黃子軍
Attend	V	V	V	V	V
Past Week Work Process					
Teammate	Work Process				
林宸宇	ASR the voice using Whisper				
張睿云	NPL splitting single text to words and simple properties				
吳姿儀	MFCC model on the voice data & splitting stop and pause on voice				

	segments	
楊書蓉	Line-Bot backend	
黃子軍	Remaining data splitting & Collecting more data	
Task Assigned		
Teammate	Working items	Temporary checkpoint
林宸宇	NPL & features seleting for text model	Simple 16 features selecting
張睿云	Voice model	MFCCs data selecting
吳姿儀	MFCCs feature selecting	MFCCs feature vector
楊書蓉	Text model	ASR+simple NPL
黃子軍	Data splitting	Dimention bank voice splitting already
Minutes Notes:		
Time	Minutes Description	
Before Meeting	ASR model training+ NPL possible barrier	
Text process in ASR	Using OpenAI developed model Whisper to transform the voice file Processed text data: <a href="https://drive.google.com/drive/folders/1x5FRmskOivtn-XY2aRzvk-SWcqxtYSbz?usp=sharing">https://drive.google.com/drive/folders/1x5FRmskOivtn-XY2aRzvk-SWcqxtYSbz?usp=sharing</a>	
Text process in NPL	Have builded the model to select simple frequency & part of the speech, but have the barrier on the same word with different meaning and meaningless words.	
Line-Bot building & APIs	Research on the ASR model Whisper embedded in Line-bot API.	
Data process in MFCCs	Using the existing model to extract the feature vector and modify the code to cut the hesitating or silent segment.	

Data splitting	Splitting the remaining data and research for the challenge data used in the previous contest,
----------------	--

Meeting Photo

