**Project Name:**

**Minutes Of Meeting (MOM) Recorder**

**(Speech to Text Convertor)**

**Project Initiators & Partners**

* Braa Mohamed Kakah (*ai.kaizenuae@gmail.com*)
* Mohamed Ahmed Al Nuaimi (*mohamedalnuaimi.1978@gmail.com*)

Table of Contents

[Introduction 3](#_Toc6771820)

[Project Flow Chat 5](#_Toc6771821)

[Coding Phase Explanation 5](#_Toc6771822)

[Difficulties and Challenges 7](#_Toc6771823)

[Future Advancements 7](#_Toc6771824)

[Resources and Links 7](#_Toc6771825)

# Introduction

This project is aim to create an application that allow the user to convert speech to organized text during a meeting. Such application will assist to record meeting minutes during the discussion live as it happened without missing any points of the discussion during the meeting.

The application shall help in record all participants discussion in the meeting and recognized the source for each voice that said from each individual. At the end, the application will convert recorded speech to organized text that presented in publishable format, in different an organized Minutes Of Meeting (MOM) as official document for any meeting established by any unit function.

To achieve the application objectives to convert speech to text from different voices, the following overall steps were considers:

* 1. Application should record the discussion during a meeting and convert them to sequence minutes.
  2. The application shall recognize different voice sources and assigned them accordingly to user name as it happened live (ex. Person 1, Person 2, Person 3… etc.)
  3. Organized the converted “Text” to a provided template of Minutes Of Meeting (MOM) in MS Word format “.docx or .doc”.
     1. Phase I will focus only to use English as language of the meeting.
     2. Phase II possibility to add additional languages (United Nation formal Languages).
     3. Phase III introduce interruption of different languages during the discussion.

Accordingly, a list of actions and tasks were defined to ensure achieving the objectives as required and resolve any challenges that face the different programming phases. Mentioned list or tasks were summarized as following points:

1. Search of any similar application that ever published to gain the lessons learned.
2. List of required Python’s libraries that required to code or program.
3. Record “Test” & “Train” voice in WAV format for any public article.
4. Train the application to recognize different voice sources and assigned them accordingly to user name as it happened live (ex. Person 1, Person 2, Person 3… etc.).
5. Converted text should be in accurate sequence.
6. Store converted “Text” to a provided template of Minutes Of Meeting (MOM) in MS Word format “.docx or .doc”.
7. Using Languages:
   1. Phase I will focus only to use English as language of the meeting.
   2. Phase II possibility to add additional languages (United Nation Formal Languages).
   3. Phase III introduce recording of interruption of different languages during the meeting.

That application will be suitable and useful to be used in wide range of business domain which will assist to accelerate publishing the Minutes of Meeting (MOM) accurate without missing any discussed points and in short due time. Not Only this, but it will help also converting any dialogue during investigation into police stations or trial sessions or even in human resources sessions, for example.

It was a challenge to complete the project in such limited period of time; however, at this moment this application need to be continue develop and completed with described phases to be a useful for business domain as it was the aim of it.

# Project Flow Chat

In the following a flow chart diagram illustrating the project workflow to achieve the final product of the coding as required.



# Coding Phase Explanation

In order to train the code, the following five paragraphs were recorded in WAV format and accordingly the recorded voice from two developers; Braa and Mohamed; were pass to the code:

1. “A massive coral reef garden the size of 60 football stadiums is to be created in the UAE - and is set to drive up eco tourism while safeguarding vital marine life.”
2. “Work is under way on the Fujairah Cultured Coral Reef Gardens, the largest project of its kind in the country, which will include the cultivation of 1.5 million coral reef colonies over the next five years and will span 300,000 square metres, helping to provide a boost to food security efforts.”
3. “The ambitious project was launched by the Ministry of Climate Change and Environment, in partnership with Fujairah Municipality, Dibba Fujairah Municipality and Fujairah Adventures and under the patronage of Sheikh Mohammed bin Hamad bin Mohammed Al Sharqi, Crown Prince of Fujairah, on Saturday.”
4. “The preservation of the country’s biodiversity is also a prime area of focus for the UAE leadership today as outlined in the UAE Vision 2021. The Ministry of Climate Change and Environment has put in place multiple programs that aim to carry out the country’s national strategy for biodiversity in collaboration with other local environmental authorities in the country.”
5. “ Ministry is keen on promoting the project as an ecotourism destination. It is also anticipated to encourage the spirit of volunteerism and community work as the cultivation of the coral reefs will depend heavily on the volunteering efforts of the youth.”

*Reference:* [*https://www.thenational.ae/uae/environment/major-coral-reef-garden-project-launched-in-fujairah-1.848464*](https://www.thenational.ae/uae/environment/major-coral-reef-garden-project-launched-in-fujairah-1.848464)

Those recorded voices were saved and labels as following:

* BTEST1.wav
* BTEST2wav
* BTEST3.wav
* BTEST4.wav
* BTEST5.wav
* MTEST1.wav
* MTEST2wav
* MTEST3.wav
* MTEST4.wav
* MTEST5.wav

Where “B” stand for Person 1 (Braa), and “M” stand for Person 2 (Mohamed).\

In the following detail description of application code:

# Difficulties and Challenges

Although the idea of the project is simple in principle; it was bit challenge in phasing and design the workflow nevertheless selecting the optimum of Python libraries and algorithm to code.

One of the main challenges that we face in this project was working in voice recognition in Time Domain; however, that been improved after applying FFT to work in Frequency Domain. Such approach enhances the outcome and give better result as it shown in the code.

Moreover, project schedule was too tight to complete the project as required and expected. In other word, the project time was challenge to deliver the project and align with the supervisor.

One of the interesting challenge or difficulty was searching for problem solving and applying new approach that was not covered in the classroom. However, as it was explained in the classroom the web pages full of articles and support blogs.

Also to come over the issue of searching for the data, the team create a solution that suite the issue by recording their voice as “Test” dataset to use in testing and learning the code. That was better solution fit to train the application to recognize the voice of two persons as start up for this application.

# Future Advancements

Currently the application was coded as supervised speech recognition due to tight schedule; hopefully it will be followed to test the code in non-supervised.

Moreover apply more than international languages and become a commercial product.

That application will be suitable and useful to be used in wide range of business domain which will assist to accelerate publishing the Minutes of Meeting (MOM) accurate without missing any discussed points and in short due time. Not Only this, but it will help also converting any dialogue during investigation into police stations or trial sessions or even in human resources sessions, for example.

# Resources and Links

Web pages:

* Spoken Speaker Identification based on Gaussian Mixture Models : Python Implementation (<https://appliedmachinelearning.blog/2017/11/14/spoken-speaker-identification-based-on-gaussian-mixture-models-python-implementation/)>
* Loading WAV Files and Showing Frequency Response (<https://pythondsp.rob-elder.com/loading-wav-files-and-showing-frequency-response/)>