The 8ᵗʰ International Competition of the Military Technical College (ICMTC 2024)

Lt. General Ibrahim Selim Award for Innovation

Details, Rules and Format

Student teams are invited to compete and demonstrate their skills and innovations during the competition held at the Military Technical College in
Kobry El-Kobba, Cairo, Egypt, on July 27ᵗʰ – August 1ˢᵗ , 2024

April 12ᵗʰ, 2024

The International Competition of the Military Technical College

Artificial Intelligence Competition (AIC - 2)

Contents

1   Overview

AIC-2 is one of the best Egypt National Artificial Intelligence Competition brought to you by The Military Technical College and The Applied Innovation Center (AIC) of the Ministry of Communications and Information Technology.

Speech technologies can enable interaction with services and applications that would be difficult to acquire and use via means other than voice. Automatic Speech Recognition (ASR) systems are revolutionizing the possibilities in this area. Recent developments in speech processing, e.g. speech recognition, speaker diarization, etc., have inspired numerous applications of speech technologies. The development of Arabic Speech Recognition has been lagging compared to the advances of first-tier languages due to many reasons such as the availability of open-source data, the limited number of involved research groups, and the wide variability existing between the formal standard Arabic and the locally spoken Arabic dialects. In this competition, we focus on Arabic Speech Recognition for Egyptian Dialects.

Many usage scenarios for ASR systems are in an outdoor context where the audio signal embeds the effects of compound noise sources. In this challenge, we target the performance of ASR in noisy environments.

By participating in this competition, you will have the opportunity to hone your skills and gain practical experience in these areas. Don't miss out on this exciting opportunity to test your skills, learn new techniques, and connect with other Artificial Intelligence enthusiasts. Register today and join us for AIC-2

Any issues not covered by these published rule sets will be addressed on a case-by-case basis by the competition board. If you have any questions and comments, kindly direct them to *competition@mtc.edu.eg*.

English is the official language of the competition, and each team must submit their deliverables in English.

2   AIC-2 Rounds

Competition will be held in two rounds: An online qualification round and an onsite (MTC site) final round. The **first round** will evaluate the performance of the submitted systems on two test sets.

- 3 Hours of Clean Egyptian Arabic Recordings.
- 3 Hours of Noisy Egyptian Arabic Recordings.

A baseline model will be given for the first round with provided performance on a development set. Moreover, a data resource consists of 100 Hours of Egyptian dialect speech data will be publicly available for the competition participants.

In the first round, the evaluation will be based on the Word Error Rate (WER) for the ASR result. The best scoring 15 teams will be qualified for the **second round** of multi-speaker ASR, as the teams will be required to provide, besides the speech recognition result, the diarization of the recording with the time segments of speakers. The provided test set will be.

- 3 Hours of multi-speaker recordings (with overlapping) of Egyptian Arabic Recordings

In the second round, the evaluation will be based on both Word Error Rate (WER) and Diarization Error Rate (DER).

Each team is expected to submit the following deliverables for **both rounds**:

- System Description Paper: A detailed document outlining the description of your system's architecture, methodologies employed, and any technical details.
- Test Set Results: Results obtained from evaluating your system on the test set.
- Model Checkpoints and Inference Scripts: Submission of model checkpoints along with corresponding inference scripts to facilitate reproducibility and assessment of your system's performance.

All participants will be provided with the same dataset, which will be made available once the competition begins. Participants are permitted to construct a system by training their Acoustic Model (AM) exclusively with the provided dataset. The utilization of any additional audio datasets beyond the provided set is prohibited. However, participants may incorporate supplementary text corpora in their training process, should the need arise.

Participants are required to develop a Speech-to-Text (STT) model starting from initial training. The practice of fine-tuning pre-existing STT models is expressly prohibited. Nevertheless, participants are permitted to utilize other pre-trained models to supplement and augment the capabilities of their overall systems.

3 Evaluation

First Round

The performance metric used in this competition is the **<u>Mean Levenshtein Distance</u>** calculated for each sentence in the char-level. **<u>Levenshtein</u>** distance between two-character sequences (predicted sequence vs. target sequence) is the minimum number of single-character edits (insertions, deletions, or substitutions) required to change one word into the other.

$$L = S + D + I$$

<u>Word Error Rate (WER)</u> is normalized Levenshtein Distance (to target sequence length), a more common metric of the performance of a speech recognition or machine translation system. The word error rate can then be computed as:

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

Where:

- "S" is the number of substitutions.
- "D" is the number of deletions.
- "I" is the number of insertions.
- "C" is the number of correct words.
- "N" is the number of words in the reference (N=S+D+C).

Second Round

In the second round, an additional metric, alongside Word Error Rate (WER), will be incorporated to assess the quality of speaker identification and timing recognition (who speaks when) within the input audio file. Specifically, we will employ the Diarization Error Rate (DER) for this purpose, defined as follows:

$$DER = \frac{False\ Alarm + Missed\ Detection + Confusion}{Reference\ Length}$$

Where:

- False Alarm: Length of segments which are considered as speech in hypothesis, but not in reference.

- Missed Detection: Length of segments that are considered as speech in reference, but not in hypothesis.

- Confusion: Length of segments that are assigned to different speakers in hypothesis and reference (after applying an optimal assignment).

- Reference Length: The total length of the reference (ground truth).
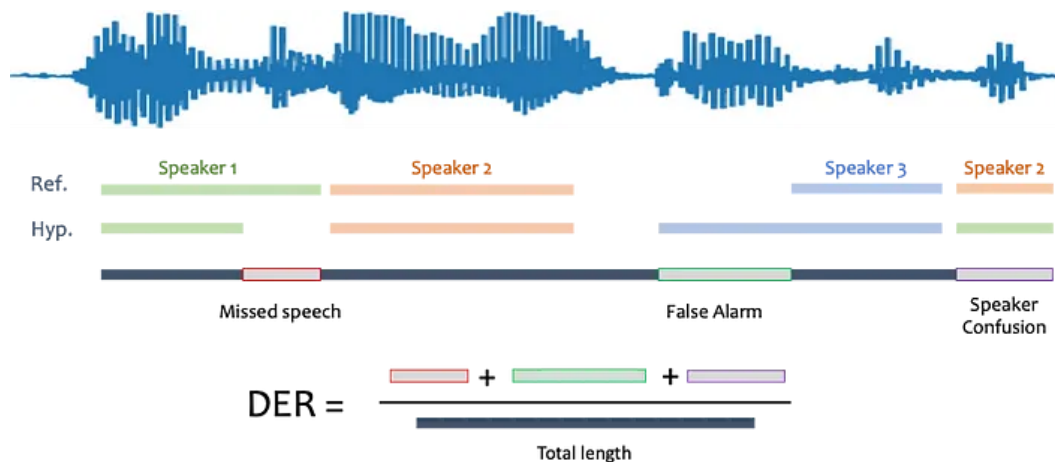
Submission Format

For the first round, **the test set (clean/noisy) audio decoding results**, the submission file should be **.csv file** and it should contain two columns: audio_id and the corresponding transcription text. Every "audio_file_id" should match one of the audio files in test set without the extension e.g. ".wav". Every corresponding transcription text should be in Arabic

Unicode format. For example,

```
audio_id,text
audioID1, السلام عليكم هل ممكن توصف لي إزاي أقدر أروح الأهرامات
audioID2, وعليكم السلام ورحمة الله الموضوع سهل تقدر تأخد مواصلة مباشرة من هنا
...
```

For the second round, **diarization test set,** the submission file should be **.rttm**. RTTM is an annotation format for audio files well designed for diarization. Diarization-type tools return one rttm per audio file, named "filename.rttm", which looks like this:

```
Audio_File_name 1 4.2 0.4 <NA> <NA> talker0 <NA>
Audio_File_name 1 5.8 1.1 <NA> <NA> talker1 <NA>
Audio_File_name 1 6.9 1.2 <NA> <NA> talker2 <NA>
Audio_File_name 1 8.1 0.7 <NA> <NA> talker1 <NA>
```

The third column indicates the onset of a region, the fourth column the duration of that region; and the identity of the speaker being indicated in the seventh column. All other columns may be ignored. Regions of non-speech are all the others (e.g., in this example, between 4.2 and 5.8). The overlap segments between speakers will be excluded from the evaluation score.

Only the qualified top 15 teams from the qualification round will be eligible to participate in the onsite final round. We encourage all Artificial Intelligence enthusiasts to participate in AIC-2, enjoy the challenges, compete for the fame and exhibit your creativity and skills to the whole world.

4   Contest Time

Qualification Round (Online): From 22nd to 28th June 2024.

Final Round (On MTC Site): From 27th July to 1st August 2024.

5   Awards

Talented Artificial Intelligence enthusiasts compete for the fame and the prize pool of 300,000 Egyptian pounds. Prizes will be awarded to the top-performing teams or individuals as the following:

➢ 1st Place: 150,000 Egyptian pounds

➢ 2nd Place: 100,000 Egyptian pounds

➢ 3rd Place: 50,000 Egyptian pounds

6   Competition Rules

6.1   Eligibility

• Undergraduate students registered in any of the Egyptian universities.

6.2   Team members

• Participants are encouraged to form teams from 3-5 members.

6.3   Registration

• Participants must register for the competition through https://aic.conferences.ekb.eg/

• Registration requires :

❖ Team Name / Members / Proof of Identity

❖ University Name / Faculty Name

❖ Proof of enrollment in the university.

❖ Official agreement letters for the participant team from their university

• Late registrations may not be accepted.

6.4   Scoring

• The team with the most points at the end of the competition will be declared the winner.

6.5   Prohibited activities

• Each participant can only join one team. Multiple accounts per user are not permitted, and neither is collaboration or membership across multiple teams.

Individuals and their submissions originating from multiple accounts will be immediately disqualified.

- Sharing of any related materials of the competition outside each individual team.

- Participants must not use any illegal or unethical method to complete challenges.

- The leaderboard will be announced in anonymous names till the organizing team verifies the final release of results.

- Violating these rules may result in disqualification.

6.6 *Disclaimers*

- The competition organizers are not responsible for any loss or damage to participants' equipment or data.

- The competition organizers reserve the right to modify the rules and format of the competition at any time.

- By participating in the competition, participants agree to abide by these rules and any decisions made by the competition organizers.