

ELC-4028 Course Project

Guidelines:

1. Each group will select just one of the project ideas below.
2. Each group will search on the web for the most cited (at least 8 references (like papers, book chapters or websites...etc.), among these references at least 4 must be very recent within the last 3 years.
3. Be careful not to copy and paste from any reference except quoted parts.
4. The expected report size is 12+ pages written in IEEE format in English (<https://essaypro.com/blog/ieee-format>).
5. A presentation of about 20+ slides is also needed. The more graphics and figures to explain the ideas the better.
6. Excellent written projects will get rewarded. Excellent project means:
 - a. The depth and the quality of the work are high.
 - b. and the written paper is good enough to be accepted for publication.
7. Each project must have at least 5 groups.
8. Deadline is Thursday 11th of May 2023. To be represented on the 18th of May 2023.

Project ideas to select from

1. The purpose of this project is to **test the level of maturity of the ChatGPT**. Do that by:
 - Use ChatGPT to solve an exam of one of the courses that you have learned, and you should know the correct answers to the questions.

Example: I asked ChatGPT this question:

A coin is thrown 3 times. What is the probability that at least one head is obtained?

I got this correct answer:

The probability of obtaining at least one head when a coin is thrown 3 times is the complement of the probability of obtaining no heads, which is the same as getting tails for all 3 tosses. Since each toss is independent, the probability of getting tails for all 3 tosses is $(1/2) * (1/2) * (1/2) = 1/8$. So, the probability of getting at least one head is $1 - 1/8 = 7/8$.

- Redo the above process with a few students who have already studied this course (those could be the team before referring to the answers). You may also ask one of your professors to help to compare the ChatGPT answers with the answers of his students.
- Compare the ChatGPT answers and the students' answers.
- You must include in your answers:
 - How do you select your test questions and why, for what course for which year...etc.
 - How do you select the students
 - Comparison of the answers
 - Add your comments and suggestions to use ChatGPT and how to lower its bad effect of there will be.

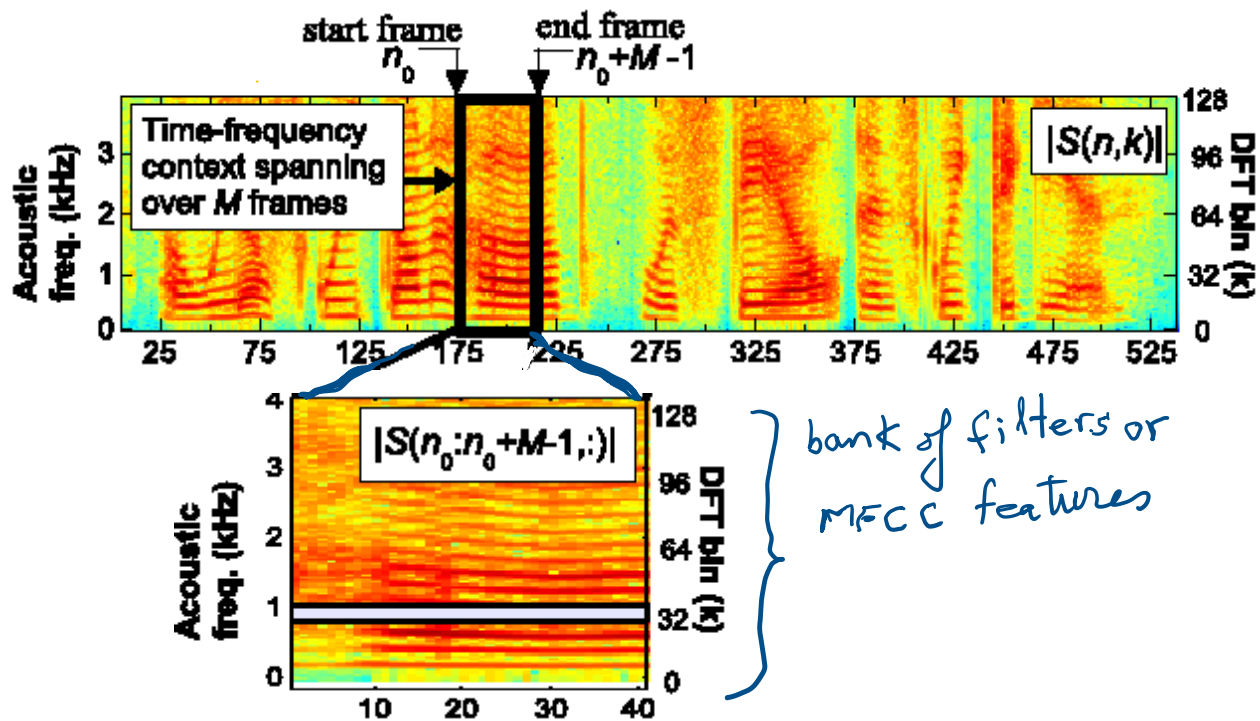
2. Select one of these Arabic technologies:

1. Automatic Speech Recognition (ASR)
 2. Typewritten Optical Character Recognition (OCR)
 3. Handwritten Optical Character Recognition (OCR)
 4. Text to Speech (TTS)
 5. Automatic Diacritization (تشكيل آلي) (AD)
 6. English to Arabic Machine Translation (MT)
- You should select at least 3-5 products from different suppliers and compare them regarding accuracy and speed and any other aspect.
 - You must design a test dataset (you may need to design it to include different genres and quality). Your selection for the test dataset should be comparable to what other research papers do. A good part of the evaluation will be on the selection and processing of the test dataset. **Do not use existing test datasets you must develop your own.**

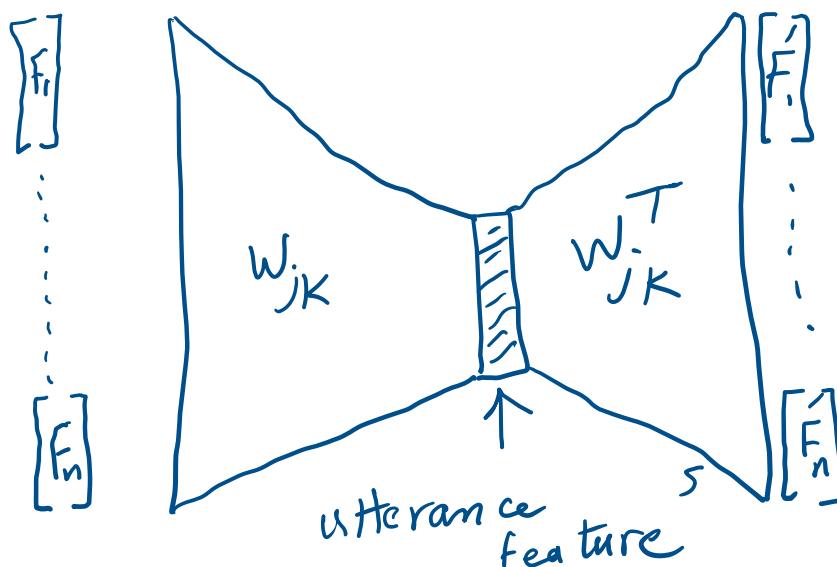
3. Autoencoder to represent and solve the utterances of the 10 digits given in assignment 2:

Given the speech data (train and test) for the 10 digits uttered by many speakers, develop an autoencoder (AE) to generate a single vector that represents each utterance (you may divide the utterance into frames of say 10 to 25 m. Seconds each, then concatenate these frames and use the AE to generate a single vector of a length of your choice to represent the complete utterance). Once you represent each utterance in a single vector, develop a classifier to classify each utterance in the test set.

1. For all the available utterances you should trim the silences before and after the utterance. Use automatic endpoint detection to do that. You may listen to the resultant utterance to be sure that the algorithm used is good enough.
2. You may start by just calculating the average frame for each utterance to get a vector that represents each utterance. Make it a baseline for your results.
3. You should try to use the AE in this way:
 - a. Use the AE to generate a vector of the same length as the 1st and 2nd frames, then between the generated vector and the 3rd frame, and so on until you finish all the utterance frames.
 - b. Repeat the algorithm in a but you may use 3 consequent frames at a time or even more. In such a case you may use zero frames to complete your last steps in some utterances.
 - c. Concatenate all the frames in one shot and generate a single vector for each utterance. However, you will have different lengths for each utterance, then use the maximum length and append zero frames in shorter utterances.



An example of a spectrum image for an utterance



Target is:

$$\min_{w_{jk}} \sum_i \|F_i - F'_i\|_2$$

An Auto Encoder is used to concatenate 2 or more frames to ultimately reduce the utterance to a single vector.