BIRZEIT UNIVERSITY

Spoken Language Processing - Spring 2024/2025

Course Project

Submission Deadline: Sunday 15/6/2025 on ITC

Project Overview:

As you know, this course requires a comprehensive term project that will constitute 15% of your course grade. Since most of these projects require an intensive workload, you can work on this project in groups of **three** (not two and not four!) students.

Project report

Each group has to submit a short report (2-4 pages long) describing their project in the IEEE transaction letter format (including appendices, figures, references, and everything else you choose to submit). The following is a suggested structure for the final report:

1. Front page (Title, student's names, ID's, etc)

2. Abstract: It should not be more than 200 words

3. Introduction: this section introduces your problem, and the overall plan for approaching your problem

4. Background/Related Work: This section discusses relevant literature for your project

5. Methodology (system description): This section details the framework of your project. Be specific, which means you might want to include equations, figures, plots, etc

6. Experiments and Results: This section begins with what kind of experiments you're doing, what kind of dataset(s) you're using, and what is the way you measure or evaluate your results. It then shows in details the results of your experiments. By details, we mean both quantitative evaluations (show numbers, figures, tables, etc) as well as qualitative results (show images, example results, etc).

7. Conclusion and future work: What have you conclude from the conducted experiments? Suggest future ideas to enhance the results.

8. References: This is absolutely necessary.

IEEE conference paper template is found on the course page at Moodle (itc.birzeit.edu).

Project Idea:

This project focuses on the pronunciation disorders of the Arabic phoneme [ر] (corresponding to [r] in the International Phonetic Alphabet (IPA)), as it is considered as one of the hardest and the most common Arabic speech disorders. Many children have problems in pronouncing /r/ sound. The /r/ sound disorder can be categorized into four classes; distortion, deletion, substituting /r/ by /gh/, substituting /r/ by /l/, in addition to the normal (correct) pronunciation.

The pronunciation disorder depends on the position of the /r/ sound in the word; at the beginning, in the middle or at the end. In this project, we focus on identifying the disorder type of /r/ at the beginning of the words.

So, in summary, you need to build a system (model) for identifying the category (class) of /r/ sound disorder when it comes at the beginning of a word.

The audio dataset is already divided into train and test subsets and both can be downloaded from the following shared folder:

https://drive.google.com/drive/folders/13PPVmCA_IbzGkzCvQxaB3NTcfUk_G6GB?usp=sharing

You can use accuracy, precision, recall and f1-score as classification performance measure.

Some good reference papers about /r/ sound disorder identification:

https://fada.birzeit.edu/bitstream/20.500.11889/4352/1/1-5_wocci2016.pdf

https://drive.google.com/file/d/1ipKD9c29YaUpCPuU0gMUzmntowdM9TDQ/view?usp=sharing

Useful tools:

- Speech Filling System (SFS): http://www.phon.ucl.ac.uk/resource/sfs/download.htm
- Praat software: http://www.fon.hum.uva.nl/praat/download_win.html
- Voicebox Matlab toolbox: http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html
- Netlab toolbox (it includes MATLAB implementation of Gaussian mixture Modelling, vector quantization, Neural networks, etc): http://www.aston.ac.uk/eas/research/groups/ncrg/resources/netlab/downloads/
- Cambridge Hidden Markov Model Toolkit (HTK): http://htk.eng.cam.ac.uk/download.shtml
- Kaldi toolkit: http://kaldi-asr.org/
- Python Google Colab: https://colab.research.google.com/?utm_source=scs-index
- Python Kaggle: https://www.kaggle.com/code