BIRZEIT UNIVERSITY

Spoken Language Processing Spring 2020

Main Course Project

Submission Deadline: Saturday 23/5/2020 (submission is only accepted on Moodle)

Project Overview:

As you know, this course requires a comprehensive term project that will constitute 50% or more of your course grade. Since most of these projects require an intensive workload, you have to work on this project in groups of **three** (maximum three!) students, or groups of two students. (group of 4 students is not accepted).

Project report

Final project reports are due at 11:55am on Tuesday, 26/5/2020. They should be submitted through Moodle.

Final project report should be 2-4 pages long 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. Title, Author(s)

2. Abstract: It should not be more than 300 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: What have you learned? Suggest future ideas.

8. References: This is absolutely necessary.

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

Project discussion:

At the end of the course, every group may be asked to provide a short presentation of their project online, followed by a short discussion. The work of every individual team member should be clarified in the discussion.

Project Idea:

In this project, you need to develop and evaluate a speaker emotion recognition system. As you know, speech signal has many information other than linguistic information such as speaker identity, speaker gender, speaker emotion state, speaker age … etc. in the project we want to develop and test an automatic system that can predict the emotional state of the speaker from his/her speech (happiness, sadness, anger, boredom, fear, neutral, etc). You can use what you have learned in the course to extract some useful features from the speech signal such as energy, zero-crossing rate, pitch frequency F0, MFCCs, LPC, … and many other feature we don't cover in the course and useful for recognizing the speaker emotional state. Such of those features are speaking rate, number of pauses, pause duration … etc.

Speech data for six emotional states is provided for this project (the emotions list above). First you may need to divide it into two subsets; training and testing. You can use this data to train your system and evaluate it with the testing dataset. You can use percentage accuracy as a system performance measure. Three days before submission deadline (i.e. on 20/5/2020), I will provide you with a set of new evaluation dataset to use your system and recognize emotional state of speakers and send me back your results. The accuracy of your system (using this evaluation data) will be part of your project grade. The three groups who will get the best accuracy will get a bonus!

The speech dataset can be also downloaded from the course page at Moodle.

More information and details about these projects will be given in the next lecture.

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/