M2FED Acoustic Subsystem Instructions

**Phase 0.** Installing required libraries.

Before anything, make sure to install all the programs in the ‘0-Library’ folder directly (using the default settings for all).

1. Microphone system drivers. If the 4-microphone system is used, install **M-Audio\_M-Track\_Quad\_2.9.64.zip.** If the 8-microphone system is used, install **Install\_M-Audio\_M-Track\_Eight\_1.0.11.zip.** You may need to connect them to the laptop before installing the software.
2. vcredist\_x64.exe; this is required for running the Acoustic main program because the program is written in C++.
3. MCR\_R2016a\_win64\_installer.exe; this is required for running the speaker identification algorithm.
4. Anaconda3-4.3.1-Windows-x86\_64.exe; this is required for running the mood detection algorithm. There are some checkboxes need to be checked.
   1. Install for **All Users**.
   2. Destination Folder should be the default value **C:\ProgramData\Anaconda3.**
   3. **Add Anaconda to the system PATH environment variable.**
   4. **Register Anaconda as the system Python 3.6.**
5. After Anaconda is installed, open Anaconda Prompt (you can type and search it from the start menu). Type the following commands. One line each time. After typing each of the following commands, the command line may say “*Proceed <[y]/n> ?”* . Type “y” and “<Enter>”. For some commands, it may take long time due to the Internet speed. If any of them hangs infinitely, close the window, and try it again. It will work eventually. Notice that the fifth command (which is italicized) is a single line command not three commands separately.

conda install python=3.5.3

conda install mingw

conda install libpython

pip install nose-parameterized

*pip install --ignore-installed --upgrade https://storage.googleapis.com/tensorflow/windows/cpu/tensorflow-1.1.0-cp35-cp35m-win\_amd64.whl*

pip install keras

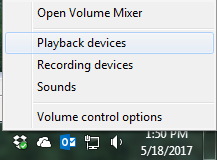
pip install genism

1. Test Anaconda. Open cmd.exe (you can type and search it from the start menu). Type “python” and “<Enter>”. If you see something includes “Python 3.5.3 |Anaconda”, type “import keras”. If you see something includes “Using TensorFlow backend.”, everything is fine. If you cannot see these lines or some errors occur, please recheck the step 4 to see if some commands failed.
2. (Optional) audacity-win-2.1.2.exe; this is optional, only needed by UVA team.

**Microphone hardware connection check. (Recommended)**

Before doing Phase 1, please follow these steps.

1. Right click the speaker icon located at the bottom-right of the screen. Choose Playback devices.



1. In the Playback tab, check the m-audio device is **NOT** be set to the default device. If it is set, choose the other speakers to be the default device. Otherwise, you will not hear any sound played from the laptop.
2. In the Recording tab, check the m-audio device (the one has ‘multichannel’ in the name) **IS** set to be the default device. If it is not, right click it, and choose Set as default device.
3. Double click the m-audio device (the one has ‘multichannel’ in the name) in the recording tab. Go to the Advance tab. Choose ‘8-channel, 24bit, 44100Hz’ in the default format setting. Click “OK”.
4. Done.

**Phase 1**. Collect speaker’s speech samples.

Objective: Speak (or read a script for more than 2 minute continuously, the longer the better but less than 5 minutes) in front of the Microphone ONE. (ATTENTION: Only speak in front of the Microphone ONE)

1. Go to the ‘Acoustic-Tools/1-Recording’ folder.
2. Be ready for the speech.
3. Run M2FEDRecorder.exe.
4. Select the correct device if required. Just enter the corresponding number and press ENTER. This step is required only at the first time you run this program. Later, it will load the previous setting. In case, you do not want to use the previous setting, delete the file ‘settings.txt’.
5. Start speaking immediately. Make sure it only records one person’s voice (in other words, everyone else should be quiet).
6. After finishing speaking, press 's' to end this program.
7. Go to the ‘Recording\_samples’ folder, and find the latest one (the one with the newest timestamp). Rename it accordingly (for example, the person ID and date. The file name does not have particular requirement. Make it understandable and well-organized for your records).
8. Do the same procedure (step 2 – step 7) for the next recording of the next person.

**Phase 2**. Train speaker models.

1. Go to the ‘Acoustic-Tools/2-Training’ folder.
2. Copy the recordings from the ‘Recording\_samples’ folder in the first step by following these steps.
   1. Go to the 'singles' folder. DO NOT REMOVE the empty folder '0-nonFamily'. (In the 'singles' folder should contain each speaker's recording sample file. For each speaker, his/her recordings should be in a separate folder. The folder name can be anything but recommend to be the speaker name/number in order to distinguish them. ) For example, it can be: Create a new folder named "1-NAME-A", and put all the recordings belong to person NAME-A in to this folder. Create a new folder named "1-NAME-B", and put all the recordings belong to person NAME-B in to this folder.
3. After copying all the recordings into their corresponding folders, check each folder and make sure one folder for one person, and the total number of folders is the same as the total number of speakers.
4. Run M2FEDTraining.exe. It should print a lot of logs on screen and finally it will create a file named 'models\_1024' in the folder. DO NOT make any changes of the file 'models\_ubm1024'.

After the Phrase 1 and Phrase 2, the system is ready for a family discussion session or real deployment.

**Phase 3**. Run the system to detect speaker id and mood in a family discussion session.

1. Go to the ‘Acoustic’ folder.
2. Copy the 'models\_1024' file from Phase 2 into the 'SpeakerIDModel' folder. DO NOT make any changes of the file 'models\_ubm1024' in the 'SpeakerIDModel' folder.
3. Run StartAcoustic.bat. It should open several windows. DO NOT close any of them, just keep them running.
4. If the session is finished, press 's' to all the windows, wait for about 30 seconds after pressing the ‘s’, and the window should show 'press any key to exit'. Now press any key, or simply close the window.

About the generated results:

If you want to see the results in real time, go to the IntegrationProgram window. It should print out speaker ID (a number), mood (a string) and microphone number.