Assignment 2a

EE698V – Machine Learning for Signal Processing

(**10 marks**)
To learn machine learning (pun unintended), people mostly use standard datasets available off the shelf. But in real life, you may work on a problem where you might (surely, if you do something original) not have one. Have you thought what kind of effort is needed to prepare a dataset of your own? What kind of challenges are involved in making sure it is actually useful for your purpose?

Let us work towards an audio classification problem. Choose some classes you want to detect in the sounds around you, say – speech, door closing/slamming, door knock, music playing, etc. Prepare a dataset of sounds by recording with any device (laptop, mobile phone, etc.). Organize the data structure so that it is easily retrievable, i.e., usable for preparing your models.

We will be using this data for many interesting tasks in this course, so prepare it carefully.

Q1. Prepare a dataset of sounds:

i. Choose a set of 5 or more categories. Include at least these:
   a. SPEECH
   b. DOORCLOSE
   c. DOORKNOCK
   d. MUSIC (playing from a device – laptop/mobile/etc.)
   e. BACKGROUND (where all the other categories are absent)

ii. Record at least 10 wav files, 10s each, of each category. Save them in a folder wav/
   a. Make sure you cover a lot of variety – different speakers, different kinds of doors, different patterns of closing and knocking, different kinds of musical instruments/genres.
   b. Make sure one recording has only one category in it. E.g. recording of speech should not contain music.

iii. Prepare a label file (labels.csv) that has wavfilename and label. An example could be:
   Rec123.wav,SPEECH
   20190911_23.wav,DOORCLOSE

   The folder structure of your submission should look like:

```
<yourRollNo>
  |
  |--- <1st recfilename.wav>
  |--- <2nd recfilename.wav>
  |       ...
  |--- labels.csv
```

Submit at: https://forms.gle/Eijj4NhndEBNK9N59

Assignment 2b

EE698V – Machine Learning for Signal Processing

- Take a printout of this pdf and write your answers here. Or else, write your answers on A4 sheets at the same locations as in this pdf.
- You can attach extra sheets at the end. But these first three sheets should contain the final answer.
- Submissions should be hand-written and handed over to your TAs before the deadline.

Q1. DSP:

a. Write your last name (surname):

b. Plot a triangular wave signal x (amplitude spanning between -1 to +1) of 0.5s duration with frequency Fs (in Hz), where Fs=length of your last name (surname). Mark the axes limits and wherever the signal is 0. (**2 Marks**)

c. Sample x at a frequency 7Hz and quantize it to 4 levels between -1 to +1 to obtain x_1. Plot the resulting time series x_1. (**4 Marks**)

d. Sample x at a frequency Fs >> Nyquist frequency to obtain a time series x_2. If you take an N-point DFT of x_2, where N>Fs, you will get the most prominent peak at k in the magnitude of the DFT? Express k in terms of N and Fs. (**4 Marks**)

Q2. Evaluation Metrics

A classifier is designed to classify rose R, lotus L and jasmine J. There are 8 test samples, on which the classifier predicts the probability of each class as follows:

P(R)	P(L)	P(J)	True label
0.37	0.63	0.	R
0.56	0.27	0.17	R
0.2	0.37	0.43	L
0.33	0.26	0.42	L
0.18	0.79	0.02	L
0.41	0.25	0.34	J
0.12	0.17	0.7	J
0.49	0.16	0.35	R

a. If the hard label is decided as $y = arg\max_{l} P(l)$,
   a. the accuracy of the classifier will be: (**1 Marks**)

   b. the confusion matrix will look like: (**2 Marks**)

   c. For lotus, the precision, recall and F1 measure will be: (**3 Marks**)

b. If I want to use this classifier to detect lotuses, I can decide the label using a threshold $\theta > 0$ as $y = L, if\ P(L) > \theta$. Draw the precision recall curve (PRC) by varying the values of θ. (**4 Marks**)

Q3. Image Filtering

An image is given as a matrix F. It is filtered with a filter matrix H.

a. Write the output G, which is a correlation of F and H, with zero padding. (**4 Marks**)
Hint: See reference: http://www.robots.ox.ac.uk/~az/lectures/ia/lect1.pdf

$$F = \begin{bmatrix} 0.6 & 0.2 & 1.2 & 0.4 & 0.3 \\ 0.9 & 1.7 & 0.3 & 1.1 & 0.1 \\ 0.6 & 0.8 & 1.1 & 0.2 & 0.5 \\ 0.4 & 0.1 & 0.4 & 0 & 0.2 \end{bmatrix} \qquad H = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

b. This operation is called as _____. (**1 Mark**)

Q4. DFT

a. Prove that for the DFT of a real time series x, just half the spectrum (i.e., k=0 to Nfft/2) is sufficient to reconstruct the entire spectrum (i.e., for any k). (**4 Marks**)
Hint: use the definition of X[k].

b. Does this hold for a complex time series too? (**1 Mark**)