Applied Machine Learning Laboratory Assignment - 3

19th Mar 2016

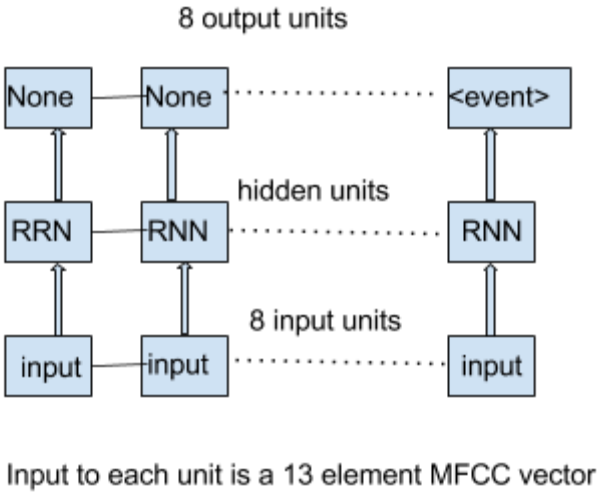
Develop a Recurrent Neural Network for classifying audio clips as one of the events - fours, sixes and wickets

You are given a dataset of audio clips pertaining to events in a cricket match with the following event space - fours, sixes and wickets.

1. The dataset given is MFCC vectors, hence, there is no need of any preprocessing / quantising the MFCC vectors. The dataset has 3 folders, namely, fours, sixes and wickets, corresponding to the respective events. Each file in the folder corresponds to MFCC vectors for an occurrence of that event.

   Instructions regarding using the dataset :
   To access the content in a pickle file,
   import pickle
   ds = pickle.load(open(<file_name>,"rb"))
   ds will now contain the data structure stored in <file_name> pickle file.

2. The input to the RNN will be a 13 element MFCC vector (size of MFCC vector).
3. Number of hidden units can be chosen according to convenience, say, 16.
4. Since each MFCC is 5 ms and it takes minimum 40 ms to make a prediction, the output is an 8 element vector.
5. The output is an 8 element vector of the following format :
   [None, None, None, None, None, None, None, <event>]



Input to each unit is a 13 element MFCC vector

6. Use 80 % of the dataset for training and 20 % of the dataset for testing.
7. Calculate accuracy (no of correct predictions / no of test samples) and report it.

Deliverables :

- Post the code on Google groups by 12:45 pm, 19th Mar 2016.
- Report the accuracy and observations of the RNN on a Facebook post by 9 pm, 20th Mar 2016.