CSL712 Assignment 2

Milan Chaudhari
2015csb1010

November 27, 2018

1 TASK 1

The task was to predict the missing data-points from the 5000 sequential data-points. For this task I experimented with an RNN and an LSTM network. Both the networks contains only one layer, which is the corresponding recurrent layer, where the last recurrent-sequence is the output of the network. The networks takes Nth to N+20th data-points and outputs the N+21st data-point.

Using sliding window method, I generated inputs and corresponding outputs from the file 'train.txt' and passed it to the keras function fit, which splits the data into training data(80%) and validation data(20%). The training and validation errors for RNN network is 0.0034 and 0.0031 respectively, while for the LSTM network as well the errors(mean squared error) are 0.0034 and 0.0031 respectively.

For the testing for the network, i.e. to predict the missing data-point, It starts with the 1st missing data-point of the missing 20 sized block, It predicts it using 20 preceding data-point available in 'train.txt', uses the predicted data-point with 19 other preceding data-points to predict the 2nd data-point of the missing block, then uses the 2 predicted data-points with 18 other preceding data-points to predict the 3rd data-point and so on. The final testing-errors for the RNN network and LSTM network are 0.038 and 0.035 respectively, which shows LSTM is good as compared to the simple RNN in this task.

2 TASK 2

The task was to predict the upcoming audio frames. The network has two layers, one recurrent layer followed by one FC layer, where recurrent layer deals with the relationship between the preceding notes and the current note, while adding an FC layer combines information from all 4 channels to predict the notes in those channels. The network takes notes from time N to N+100 in all 4 channels and predict notes at time N+101. The network with RNN layer gives 0.0440 training and 0.0968 validation error(mean squared error), while the LSTM network gives 0.0437 training and 0.1034 validation error. Here LSTM

shows less training error, but goes under a bit of over-fitting as compared to RNN.

For the given sequence of notes for 4 channels for time T, the notes at T+1 time are predicted, using which along with the given sequence of notes the notes at T+2 time are predicted and so on till T+100. For the RNN and LSTM networks, the output .txt file are 'out_srnn.txt' and 'out_lstm.txt', while the audio output files are 'out_srnn.wav' and 'out_lstm.wav' respectively.