

Grapheme-Phoneme Conversion using Neural Networks

M.S.Krishna Deepak Jagadeesh Nelaturu Sashank Gondala

IIT Bombay

April 28, 2015

- We try to solve the grapheme-phoneme conversion problem using feed forward n/w trained using backpropagation.
- Encoding is done in the following way. Each phoneme is represented by 7 bits as we have 69 different phonemes. Each grapheme is represented by 5 bits as we have 26 different graphemes - letters A to Z.
- The neural networks has three layers. The input layer, a hidden layer and an output layer.
- The number of neurons in the hidden layer is chosen to be the average of number of neurons in input layer and hidden layer.

Grapheme-Phoneme conversion

- In grapheme to phoneme conversion, size of input layer is $5*L$, size of output layer is $7*L$ and hence size of hidden layer is $6*L$.
- We set the initial weights randomly and train this neural network using training data.
- Backpropagation formulae are used to generate new weights after reading each line of the corpus.
- Accuracy is measured by seeing how many graphemes have given correct phonemes on a testing data set. We have achieved an accuracy of around 30 percent.

Phoneme-Grapheme conversion

- In phoneme to grapheme conversion, size of input layer is $7*L$, size of output layer is $5*L$ and hence size of hidden layer is $6*L$.
- We train the neural network in a similar way and use backpropagation formulae to find the weights in neural network.
- Accuracy is measured similarly by seeing how many phonemes have given correct graphemes on a testing data set. We have achieved an accuracy of around 60 percent.