Working of Sentimental Analysis on Open-Ended Conversations using LSTM

1. Function Definitions

❖**Clean_data**(text, lemmatize = True)

- ✓ Performs **pre-processing on Input sequences.**
- ✓ Remove HTML tags
- ✓ Fix mappings
- ✓ Clean emojis
- ✓ Remove urls
- ✓ Remove Punctuations
- ✓ Perform lemmatization

❖**get_corpus**(text)

- ✓ Returns the **list of separate words** in the text.

❖**get_ngrams**(review,n,g)

- ✓ Returns the **frequency-wise sorted list of n-grams**. (g is given for n-gram range (g to g)).

❖ *Model Building:*

➢ **LSTM** type:

- **INPUT Layer**: Shape = Max. length of input.
- **EMBEDDING**: Perform word embedding with Vocabulary size = V+1, Output = 5, Input length = Max. length of input.
- **Batch Norm**: Performed on the layers to normalize the weighted sum of every neuron.
- **DropOut**: To Spread out weights on next layer.
- **Conv1D**: Perform convolution over one direction with stride 1 with 'RELU' activation function.
- **DropOut**: To Spread out weights on next layer.
- **Max Polling**: To extract dominating features.
- **LSTM**: Implementing LSTM by taking output dimension as 128.
- **LSTM**: Implementing LSTM by taking output dimension as 64.
- **DropOut**: To Spread out weights on next layer.
- **Dense**: Single output neuron layer with Sigmoid activation.

✓ Model is trained on **5 epochs** using **'Adam'** optimizer.
✓ Loss is measured on **Binary crossentropy**.

❖ ***Accuracy:***

The accuracy of test results was found to be **88.32%.** The analysis of the result is mentioned below: