Assignment No. 5

Aim: Implement the Continuous Bag of Words (CBOW) Model. Stages can be:

a. Data preparation

b. Generate training data

c. Train model

d. Output

Objective: To learn and understand continuous bag of words model.

Infrastructure: Computer/ Laptop/ Virtual Machine

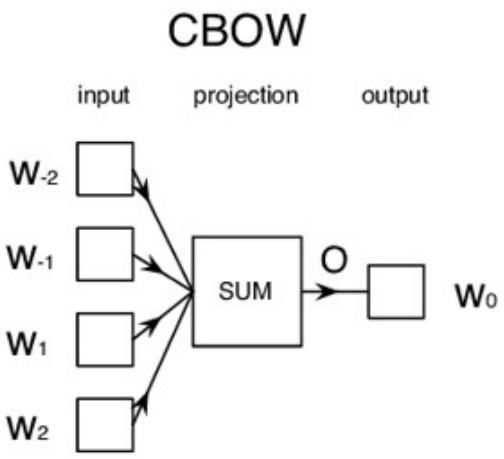
Software used: Jupyter Notebook/Google Colab, Tensorflow, Kearas

Theory:

What is the CBOW Model?

The CBOW model tries to understand the context of the words and takes this as input. It then tries to predict words that are contextually accurate. Let us consider an example for understanding this. Consider the sentence: 'It is a pleasant day' and the word 'pleasant' goes as input to the neural network. We are trying to predict the word 'day' here. We will use the one-hot encoding for the input words and measure the error rates with the one-hot encoded target word. Doing this will help us predict the output based on the word with least error.

The Model Architecture



The CBOW model architecture is as shown above. The model tries to predict the target word by trying to understand the context of the surrounding words. Consider the same

sentence as above, 'It is a pleasant day'.The model converts this sentence into word pairs in the form (contextword, targetword). The user will have to set the window size. If the window for the context word is 2 then the word pairs would look like this: ([it, a], is), ([is, pleasant], a),([a, day], pleasant). With these word pairs, the model tries to predict the target word considered the context words.

If we have 4 context words used for predicting one target word the input layer will be in the form of four 1XW input vectors. These input vectors will be passed to the hidden layer where it is multiplied by a WXN matrix. Finally, the 1XN output from the hidden layer enters the sum layer where an element-wise summation is performed on the vectors before a final activation is performed and the output is obtained.

Implementation of the CBOW Model

- Import the libraries and read our dataset.
- For the implementation of this model, we will use a sample text data

- Generate function that create window sizes and pairs of target words
- Build neural network on sample data

Conclusion

we saw what a CBOW model is and how it works. These can be used for text recognition, speech to text conversion etc.