1. My first step is to look at the data provided if there is need to do preprocessing or not. By looking at the data, there are no commas and regular expressions so I didn’t need to do that. Then I checked if there were any missing values or not. And I found some missing rows in the given data, so I removed them and reset the index.

2. Now as it can be seen from the data that rows for a particular class are together. So before splitting the data between the train set and test set, I shuffled the data so that no classes could be missed during the training of our model.

3. Next step is to check the number of unique classes in the dataset, which was 11 for our case. Then using tokenizer (maximum word taken - 20000), I found out the bags of words and after that I found out the frequency of each word and then map each of them to a number in their decreasing order of frequency. As we can see from the plot, there were 40-60 words which had very few counts, so I neglected them and used the rest of the words for training our model.

4. The length of the sentences are different, so the average number of words taken in a sentence is 40. So if there is less number of words in a sentence, then I have added 0 to make the length 40.
5. The splitting was done in the ratio of 80% and 20% for training and validation sets respectively. Finally, the model was fit to the training data and accuracies were calculated for the training and validation set.
6. As we can see that the validation set accuracy (0.6448) is less than that of the training accuracy(0.8589). That means that our model is overfitting the training dataset. To reduce the overfitting, tuning parameters can be changed accordingly.