# OUTPUT

## Output Explanation:-

The output screenshot attached below is taken after 1 fold of 3 fold cross validation of 2nd order HMM with viterbi algorithm for the prediction of the best Part of speech tags.

**Important facts:-**
- The processing time for one fold of 3 cross fold validation for 18K sentences as test set with 2nd order HMM with no tag reductions took >12 Hours making the complete execution of 3 cross fold validation >36 Hours.

## Output Screenshot:-

```
Select root@LAPTOP-EH5BDKFN: /mnt/c/Users/pruty/Desktop/NLP_assignment 2

root@LAPTOP-EH5BDKFN:/mnt/c/Users/pruty/Desktop/NLP_assignment 2# python3.5 code.py

----------File Read Successfully---------


-------------FOLD 1---------------

Traning HMM model..............
Model Training successful

-----------Training Stats----------
Number of sentences -  34563
Total number of words -  773064
Total number of tags -  773064
Unique words -  46615
Unique tags -  430
Maximum sentence length -  221
Minimum sentence length -  3



------2nd order Hidden Markov Model with Vitirbi algorithm------

Accuracy =  0.7666397453560739

F1_score of NN, JJ, VBD =  [0.7833274709813578, 0.7338638373121131, 0.7492447129909364]
Average F1 score over all tags =  0.6878797067091328
Number of F1_scores =  430

precision of NN, JJ, VBD =  [0.7577407281388228, 0.7136715391229579, 0.7366336633663366]
Average precision over all tags =  0.7657329741081819
Number of precisions =  430

Recall of NN, JJ, VBD = [0.8107025846377867, 0.7552320291173794, 0.7622950819672131]
Average recall over all tags =  0.6798209751542802
Number of recalls =  430

Confusion matrix dimensions (row, column) : [430, 430]
Subset of confusion matrix:-
        NN      IN      JJ      NNS     AT
NN      2227    85      40      57      5
IN      77      1903    24      32      85
JJ      97      9       830     34      0
NNS     9       16      1       747     1
AT      135     4       64      31      1578

Top 5 most incorrectly predicted word types = ['NN', 'NNS', 'IN', 'JJ', '.']
```

## Output Statistics:-

Accuracy for **Fold 1 -** 76.66%
Average **Precision -**   76.57%
Average **F1-score** - 68.78%
Average **Recall** - 67.98%
Top 5 **most incorrectly predicted** word types - ["NN", "NNS", "IN", "JJ", "."]

Since printing out all the statistics was not feasible, therefore I've predicted few values from each statistic and the dimension of the total statistic matrix which you can verify from my code that all the values for all the tags have been computed.