Name : Ammaar Ahmad, Ayush Pandey
Roll No : 1801CS08, 1801CS11
Team Code : 1801cs08_1801cs11
CS563 NLP Assignment 2

Codalab Link :
https://colab.research.google.com/drive/1OMcpa2vhukSdFv7k8eGozDLQPKoAzdIw?usp=sharing
Output Files - HMM_21_results.csv, HMM_3_results.csv

HMM for 3 tags (BIO) Output

```
------------------------------------------------------------------------
------
HMM for 3 tags :


Trigram Model with No Context Validation Results
Evaluating 1000 sentences.

100%|██████████| 1000/1000 [00:02<00:00, 496.72it/s]
HMM Model Accuracy = 0.9037574564909907

Class-wise Accuracies

| Class (Tag)    |   Accuracy |
|----------------+------------|
| B              |   0.20121  |
| I              |   0.117773 |
| O              |   0.9587   |


Bigram Model with No Context Validation Results
Evaluating 1000 sentences.

100%|██████████| 1000/1000 [00:00<00:00, 2255.55it/s]
HMM Model Accuracy = 0.9015435705061189

Class-wise Accuracies

| Class (Tag)    |   Accuracy |
|----------------+------------|
| B              |   0.2118   |
| I              |   0.100642 |
```

| O | 0.956387 |

Trigram Model with Context Validation Results
Evaluating 1000 sentences.

100%|██████████| 1000/1000 [00:02<00:00, 444.58it/s]
HMM Model Accuracy = 0.912121025767173

Class-wise Accuracies

Class (Tag)	Accuracy
B	0.192133
I	0.130621
O	0.967687

Bigram Model with Context Validation Results
Evaluating 1000 sentences.

100%|██████████| 1000/1000 [00:00<00:00, 2066.48it/s]
HMM Model Accuracy = 0.9114445606051288

Class-wise Accuracies

Class (Tag)	Accuracy
B	0.193646
I	0.122056
O	0.967158

Best Model Test Dataset Results

Evaluating 3849 sentences.

100%|██████████| 3849/3849 [00:08<00:00, 443.15it/s]HMM Model Accuracy =
0.8859377019516609

Class-wise Accuracies

Class (Tag)	Accuracy
B	0.183703
I	0.130137
O	0.963068

HMM for 21 tags (BIO - NER) Output

```
--------------------------------------------------------------------------------
------
HMM for 21 tags :


Trigram Model with No Context Validation Results
Evaluating 1000 sentences.

100%|██████████| 1000/1000 [02:39<00:00,  6.28it/s]
HMM Model Accuracy = 0.9002521370149438

Class-wise Accuracies

| Class (Tag)    |   Accuracy |
|----------------+------------|
| B-company      |  0.153846  |
| B-facility     |  0.0526316 |
| B-loc          |  0.284483  |
| B-movie        |  0         |
| B-musicartist  |  0.0243902 |
| B-other        |  0.106061  |
| B-person       |  0.157895  |
| B-product      |  0.189189  |
| B-sportsteam   |  0.0285714 |
| B-tvshow       |  0         |
| I-company      |  0         |
| I-facility     |  0.102564  |
| I-loc          |  0.166667  |
| I-movie        |  0         |
| I-musicartist  |  0.0285714 |
| I-other        |  0.0412371 |
| I-person       |  0.0842105 |
| I-pr90.942oduct     |  0.0165289 |
| I-sportsteam   |  0.0769231 |
| O              |  0.959492  |


Bigram Model with No Context Validation Results
Evaluating 1000 sentences.

100%|██████████| 1000/1000 [00:06<00:00, 154.08it/s]
HMM Model Accuracy = 0.8979152573642457

Class-wise Accuracies
```

Class (Tag)	Accuracy
B-company	0.153846
B-facility	0.0526316
B-loc	0.284483
B-movie	0
B-musicartist	0.0243902
B-other	0.106061
B-person	0.175439
B-product	0.189189
B-sportsteam	0.0285714
B-tvshow	0
I-company	0
I-facility	0.102564
I-loc	0.166667
I-movie	0
I-musicartist	0.0285714
I-other	0.0412371
I-person	0.0631579
I-product	0.00826446
I-sportsteam	0.0769231
O	0.956981

Trigram Model with Context Validation Results
Evaluating 1000 sentences.

100%|██████████| 1000/1000 [02:50<00:00, 5.86it/s]
HMM Model Accuracy = 0.9095381587848226

Class-wise Accuracies

Class (Tag)	Accuracy
B-company	0.153846
B-facility	0.0526316
B-loc	0.25
B-movie	0
B-musicartist	0.0243902
B-other	0.0833333
B-person	0.128655
B-product	0.108108
B-sportsteam	0
B-tvshow	0
I-company	0
I-facility	0.0769231
I-loc	0.166667

I-movie	0
I-musicartist	0
I-other	0.0412371
I-person	0.136842
I-product	0.00826446
I-sportsteam	0
O	0.970528

Bigram Model with Context Validation Results
Evaluating 1000 sentences.

100%|██████████| 1000/1000 [00:09<00:00, 107.74it/s]
HMM Model Accuracy = 0.9094151651189963

Class-wise Accuracies

Class (Tag)	Accuracy
B-company	0.153846
B-facility	0.0263158
B-loc	0.25
B-movie	0
B-musicartist	0.0243902
B-other	0.0833333
B-person	0.140351
B-product	0.108108
B-sportsteam	0
B-tvshow	0
I-company	0
I-facility	0.0512821
I-loc	0.190476
I-movie	0
I-musicartist	0
I-other	0.0515464
I-person	0.147368
I-product	0
I-sportsteam	0
O	0.970264

Best Model Test Dataset Results

Evaluating 3849 sentences.

100%|██████████| 3849/3849 [11:26<00:00, 5.61it/s]
HMM Model Accuracy = 0.8806061781052087

```
Class-wise Accuracies

| Class (Tag)    |   Accuracy |
|----------------+------------|
| B-company      |  0.0772947 |
| B-facility     |  0.0632411 |
| B-loc          |  0.226757  |
| B-movie        |  0         |
| B-musicartist  |  0.0052356 |
| B-other        |  0.0273973 |
| B-person       |  0.136929  |
| B-product      |  0.0203252 |
| B-sportsteam   |  0.0204082 |
| B-tvshow       |  0         |
| I-company      |  0.045283  |
| I-facility     |  0.103825  |
| I-loc          |  0.0913242 |
| I-movie        |  0         |
| I-musicartist  |  0.0142857 |
| I-other        |  0.0647482 |
| I-person       |  0.103333  |
| I-product      |  0.002     |
| I-sportsteam   |  0         |
| I-tvshow       |  0         |
| O              |  0.965499  |
--------------------------------------------------------------------------------
------
```

Validation Accuracy

HMM Model	**Emission Probability Context**	**Accuracy (BIO)**	**Accuracy (BIO-NER)**
Trigram	No	90.375	90.025
Bigram	No	90.154	89.792
Trigram	Yes	91.212	90.954
Bigram	Yes	91.114	90.942

Test Results

For both BIO labels with and without NER performs best with Trigram model with context information in emission probability

BIO Tags (3 labels only) Accuracy = 88.59%
BIO NER Tags (21 labels) Accuracy = 88.06%

We achieve better accuracy for the 3-tag model as compared to 21-tag model. Increasing number of tags demands finer distinction among tags which lead to lower accuracy. Also -tag tag consumes less time as time complexity of viterbi algorithm is O(P*P*L) where P is number of tags