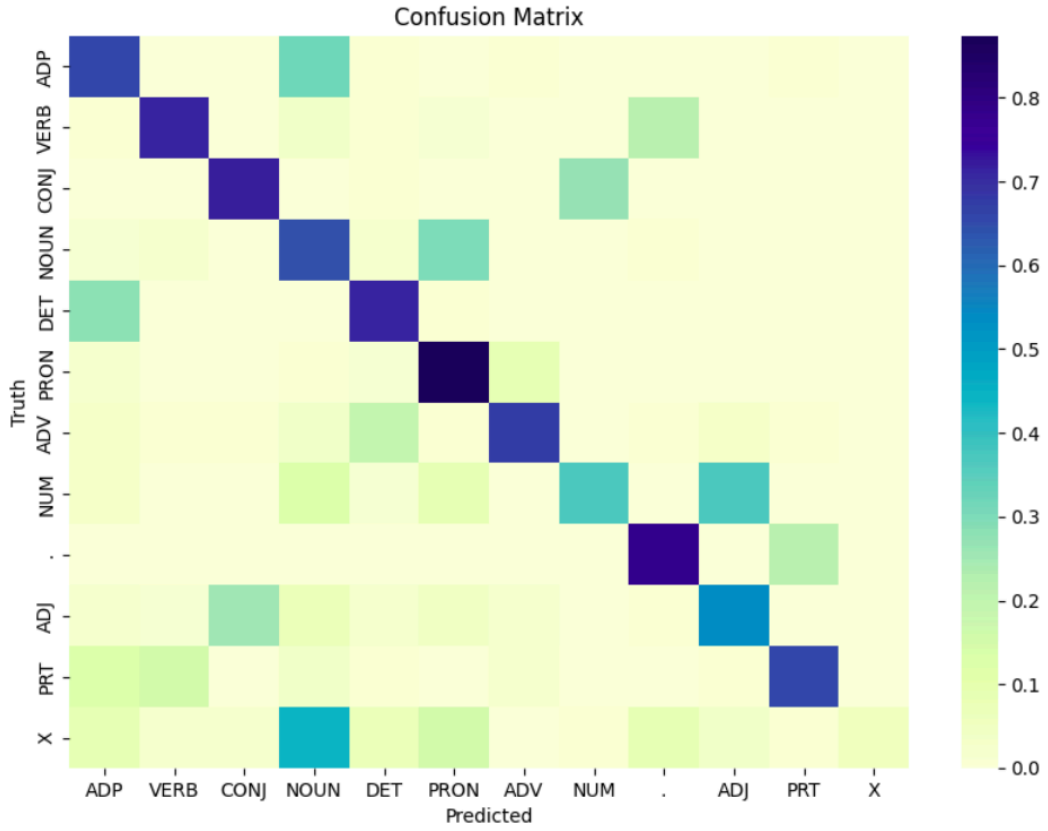
Rupak Biswas
2001CS57
NLP-Assignment

Confusion matrix of HMM (Bigram)



Overall accuracy : 95.45%

5 for cross validation results:

State	Accuracy	Precision	Recall	F1 Score
ADP	0.983957	0.881899	0.974557	0.925916
VERB	0.982741	0.966669	0.936158	0.951169
CONJ	0.999556	0.993949	0.991698	0.992822
NOUN	0.969981	0.901290	0.946194	0.923197
DET	0.990395	0.937608	0.982320	0.959443
PRON	0.996144	0.987544	0.960845	0.974012

```
ADV         0.987918   0.933304   0.852360   0.890997
NUM         0.998130   1.000000   0.725557   0.840954
.           0.999474   0.996570   1.000000   0.998282
ADJ         0.986691   0.915053   0.813964   0.861553
PRT         0.990150   0.927287   0.787711   0.851819
X           0.999322   0.833333   0.080645   0.147059
Worker Process-5 done. Total Accurecy : 0.9422294928721664


State       Accuracy   Precision  Recall           F1 Score
ADP         0.983254   0.875947   0.977241   0.923826
VERB        0.984279   0.969795   0.943220   0.956323
CONJ        0.999606   0.993402   0.993766   0.993584
NOUN        0.971249   0.909425   0.946184   0.927441
DET         0.990304   0.936178   0.981312   0.958214
PRON        0.996576   0.984346   0.969027   0.976626
ADV         0.988851   0.933604   0.865598   0.898316
NUM         0.997872   0.995976   0.725806   0.839695
.           0.999640   0.997536   1.000000   0.998766
ADJ         0.987004   0.917856   0.816077   0.863979
PRT         0.989729   0.947748   0.783907   0.858077
X           0.999065   0.666667   0.069767   0.126316
Worker Process-4 done. Total Accurecy : 0.9437136807964323


State       Accuracy   Precision  Recall           F1 Score
ADP         0.984312   0.886334   0.969964   0.926265
VERB        0.984480   0.969837   0.942328   0.955885
CONJ        0.999642   0.994016   0.995803   0.994909
NOUN        0.970728   0.903210   0.944520   0.923403
DET         0.989699   0.930827   0.976052   0.952903
PRON        0.996181   0.981031   0.969438   0.975200
ADV         0.988679   0.921897   0.873439   0.897014
NUM         0.998906   0.989873   0.796334   0.882619
.           0.999474   0.996651   1.000000   0.998323
ADJ         0.984828   0.916974   0.811797   0.861186
PRT         0.990383   0.940890   0.784348   0.855517
X           0.998632   0.583333   0.053030   0.097222
```

Worker Process-6 done. Total Accurecy : 0.942971980513673

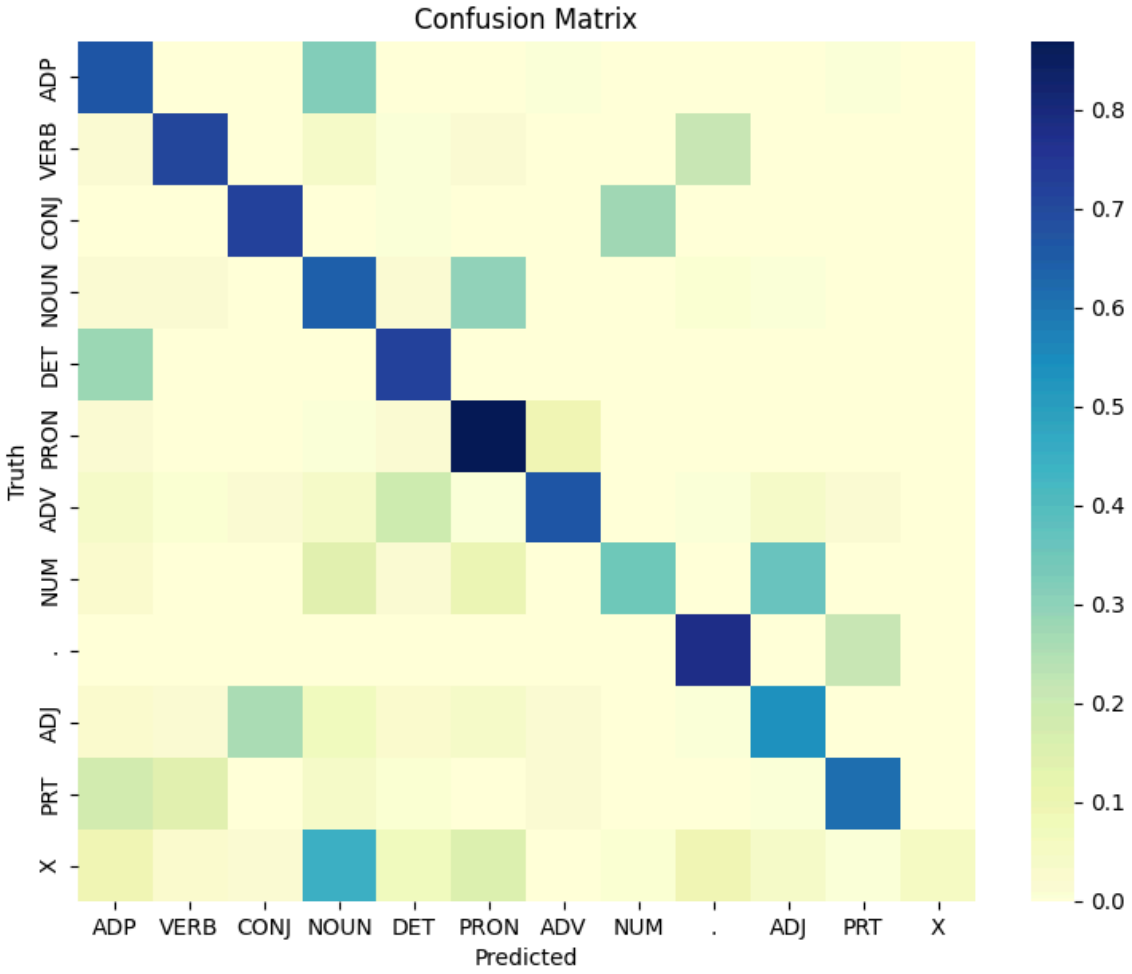
State	Accuracy	Precision	Recall	F1 Score
ADP	0.985916	0.917384	0.980641	0.947958
VERB	0.979155	0.947344	0.919067	0.932992
CONJ	0.999523	0.996608	0.988702	0.992639
NOUN	0.962617	0.908280	0.940959	0.924330
DET	0.987057	0.909273	0.992403	0.949021
PRON	0.997640	0.975987	0.953321	0.964521
ADV	0.989699	0.934741	0.851601	0.891236
NUM	0.995897	0.994700	0.684498	0.810947
.	0.998695	0.989094	1.000000	0.994517
ADJ	0.981672	0.947609	0.805105	0.870564
PRT	0.994795	0.953583	0.809261	0.875514
X	0.998921	0.818182	0.062069	0.115385

Worker Process-3 done. Total Accurecy : 0.9357937270901232

State	Accuracy	Precision	Recall	F1 Score
DET	0.984680	0.889733	0.996005	0.939875
PRT	0.996282	0.941856	0.850652	0.893933
ADJ	0.977202	0.954219	0.766894	0.850362
ADP	0.985915	0.930290	0.977467	0.953295
ADV	0.990882	0.922867	0.848165	0.883940
NOUN	0.958384	0.908120	0.936828	0.922251
PRON	0.998216	0.955083	0.943191	0.949099
CONJ	0.999458	0.995891	0.987974	0.991917
VERB	0.979583	0.937664	0.919654	0.928571
NUM	0.994388	0.997963	0.750843	0.856943
.	0.998566	0.986818	0.999871	0.993301
X	0.998861	0.444444	0.024242	0.045977

Worker Process-2 done. Total Accurecy : 0.9312093083883892

Confusion Matrix of HMM (Trigram)



Overall accuracy : 93.37%

5 fold cross validation result:

State	Accuracy	Precision	Recall	F1 Score
ADP	0.980498	0.856073	0.974216	0.911332
VERB	0.979785	0.958754	0.927307	0.942768
CONJ	0.999533	0.992825	0.992075	0.992450
NOUN	0.967422	0.897054	0.936634	0.916417
DET	0.990442	0.935116	0.985755	0.959768
PRON	0.996658	0.997573	0.957893	0.977330
ADV	0.987053	0.936310	0.833199	0.881750
NUM	0.997885	1.000000	0.689537	0.816244
.	0.998808	0.992259	1.000000	0.996115
ADJ	0.986714	0.910646	0.819247	0.862532

```
PRT         0.987415   0.902213   0.728869   0.806330
X           0.999322   0.833333   0.080645   0.147059
Worker Process-5 done. Total Accurecy : 0.9357677027342837


State       Accuracy   Precision  Recall          F1 Score
ADP         0.979943   0.851160   0.977999   0.910182
VERB        0.980889   0.958123   0.936185   0.947027
CONJ        0.999595   0.992316   0.994499   0.993407
NOUN        0.968636   0.907135   0.934122   0.920431
DET         0.990315   0.932575   0.985785   0.958442
PRON        0.996926   0.993246   0.964907   0.978872
ADV         0.988051   0.934276   0.849762   0.890018
NUM         0.997680   0.995833   0.700880   0.822719
.           0.998964   0.992947   1.000000   0.996461
ADJ         0.987094   0.913883   0.822311   0.865682
PRT         0.986689   0.932247   0.715951   0.809907
X           0.999065   0.666667   0.069767   0.126316
Worker Process-4 done. Total Accurecy : 0.9369228343619082


State       Accuracy   Precision  Recall          F1 Score
ADP         0.980714   0.858544   0.969964   0.910859
VERB        0.980987   0.960097   0.932185   0.945935
CONJ        0.999600   0.991945   0.996703   0.994318
NOUN        0.967982   0.899777   0.932466   0.915830
DET         0.989562   0.924905   0.981965   0.952581
PRON        0.996517   0.994792   0.960065   0.977120
ADV         0.987690   0.922798   0.853308   0.886694
NUM         0.998885   0.989822   0.792261   0.880090
.           0.998527   0.990680   1.000000   0.995318
ADJ         0.984449   0.907108   0.815245   0.858727
PRT         0.987774   0.923704   0.722899   0.811057
X           0.998632   0.583333   0.053030   0.097222
Worker Process-6 done. Total Accurecy : 0.935659347018655


State       Accuracy   Precision  Recall          F1 Score
```

```
ADP         0.984963   0.908629   0.983987   0.944808
VERB        0.978405   0.954590   0.906346   0.929843
CONJ        0.999523   0.996608   0.988702   0.992639
NOUN        0.961422   0.902733   0.942569   0.922221
DET         0.987784   0.914191   0.992532   0.951752
PRON        0.997812   0.983662   0.950766   0.966934
ADV         0.989058   0.932586   0.839931   0.883837
NUM         0.995522   0.994465   0.655319   0.790033
.           0.998015   0.983505   1.000000   0.991684
ADJ         0.981039   0.935366   0.808167   0.867127
PRT         0.993966   0.940981   0.782308   0.854340
X           0.998921   0.818182   0.062069   0.115385
Worker Process-3 done. Total Accurecy : 0.9332145341430436
```

```
State       Accuracy   Precision  Recall           F1 Score
DET         0.983555   0.882317   0.996062   0.935746
PRT         0.995177   0.927524   0.800745   0.859484
ADJ         0.976475   0.943042   0.767869   0.846488
ADP         0.985414   0.927588   0.977094   0.951697
ADV         0.990128   0.909107   0.843137   0.874880
NOUN        0.957245   0.903986   0.937271   0.920327
PRON        0.998491   0.974556   0.938911   0.956401
CONJ        0.999335   0.994651   0.985528   0.990069
VERB        0.978575   0.944684   0.904488   0.924149
NUM         0.993894   0.997902   0.728777   0.842366
.           0.997043   0.975810   0.996903   0.986244
X           0.998827   0.285714   0.024242   0.044693
Worker Process-2 done. Total Accuracy : 0.9270792598843312
```

Taken Assumptions:

If a work/token is not found in the calculated emission probabilities the global emission probabilities are used, in case of transition probability , 0 is the default value for both bigram and trigram model. Only difference between bigram and

trigram models is the transition probability. Prediction is done using the Viterbi algorithm in both cases.

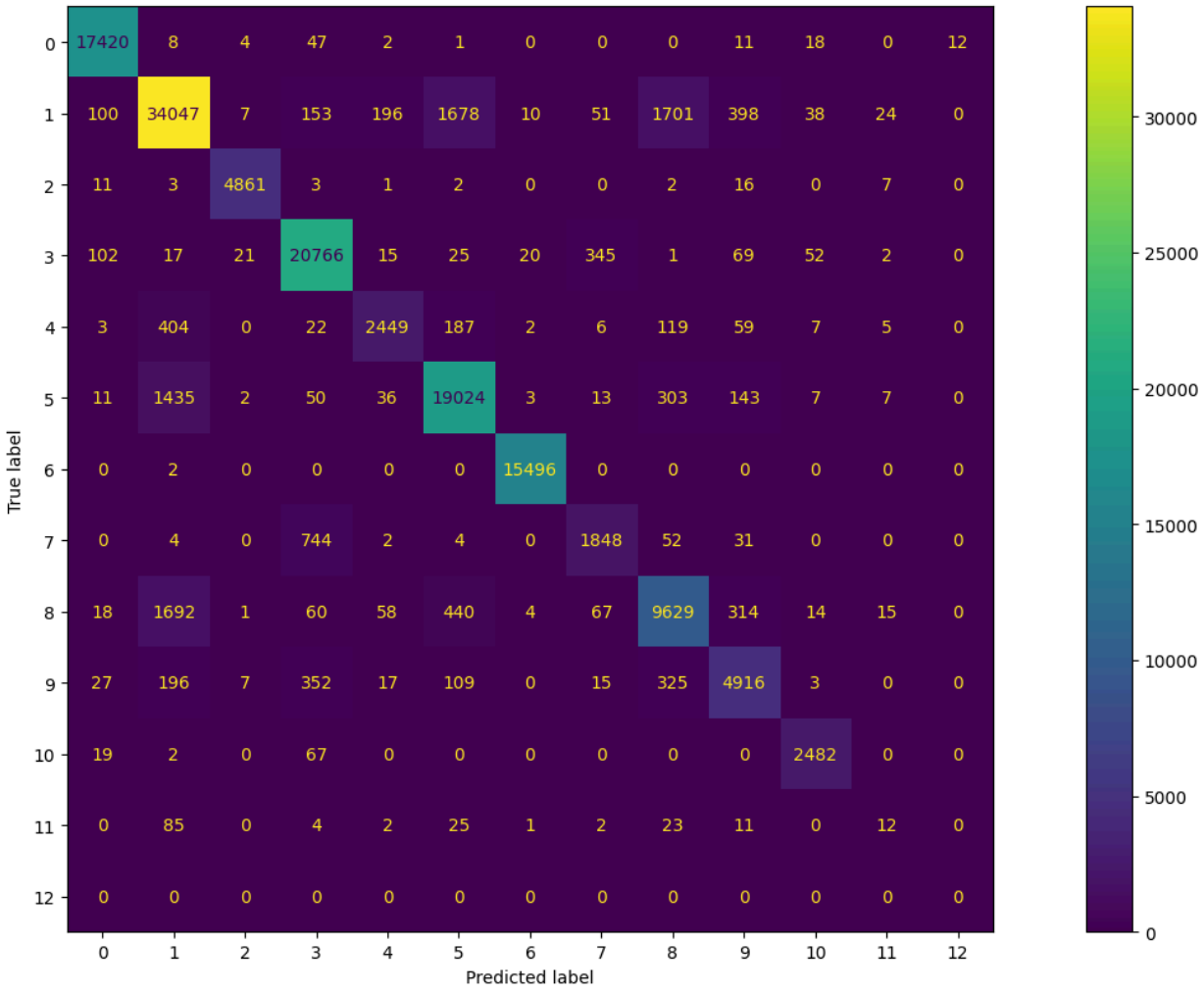
Good Example:

Bigram

Included/VERB in/ADP the/DET findings/NOUN are/VERB :/. 1/NUM ./.
['VERB', 'ADP', 'DET', 'NOUN', 'VERB', '.', 'NUM', '.'] <labels>
['VERB', 'ADP', 'DET', 'NOUN', 'VERB', '.', 'NUM', '.'] <predt>

Reasoning: The word are connected , have good correlation

Confusion matrix RNN



Accuracy : 91.2 % @ 10 epoch

RNN vs HMM

HMMs offer simplicity and ease of implementation, making them suitable for smaller datasets with limited ambiguity and real-time applications requiring low latency. Conversely, RNNs, while more complex and computationally demanding, excel with large datasets, ambiguous contexts, and the need for high accuracy, owing to their ability to capture longer-range

dependencies and contextual nuances. However, both models have their limitations; HMMs struggle with ambiguous words and unseen data, while RNNs may face challenges with overfitting, vanishing/exploding gradients, and learning long-range dependencies. A thoughtful selection between the two depends on factors such as data size, desired accuracy, and computational resources, with potential strategies including hybrid models, ensemble methods, and task-specific adaptations to optimize performance.