**REPORT - ASSIGNMENT 4**

* Feature Description :
  + Lowercasing (previous, current and next words)
  + Checking for uppercase
  + Checking for title
  + Checking for digit
  + Checking for later part of words(last 3/2 letters)
* Hyperparameters tuned (if any) :
  + Algorithm = ‘lbfgs’
  + C1 = 0.1,
  + C2 = 0.1

|  |  |  |  |
| --- | --- | --- | --- |
| # |  | Train | Test |
| 1. | 10 most common transition features | VERB -> AUX 3.844053  PROPN -> PROPN 2.769828  ADJ -> NOUN 2.235242  AUX -> AUX 2.221800  AUX -> SCONJ 2.138434  PROPN -> ADP 2.065055  VERB -> SCONJ 2.061378  NUM -> NOUN 1.749800  PROPN -> CCONJ 1.654623  DET -> ADJ 1.643873 |  |
| 2. | 10 least common transition features | AUX -> VERB -1.179604  PROPN -> AUX -1.195544  ADV -> PROPN -1.219184  AUX -> ADJ -1.222223  ADV -> AUX -1.293907  CCONJ -> AUX -1.543189  AUX -> ADP -1.613465  ADJ -> PRON -1.816283  DET -> ADP -2.150266  ADJ -> ADP -2.205392 |  |
| 3. | Precision (per tag) | PART -> 1.000  CCONJ -> 0.993  SCONJ -> 0.984  ADJ -> 0.998  ADP -> 0.998  ADV -> 0.982  VERB -> 0.998  DET -> 0.996  NOUN -> 0.998  PRON -> 0.998  PROPN -> 1.000  NUM -> 1.000  PUNCT -> 1.000  AUX -> 0.984 | PART -> 0.909  CCONJ -> 1.000  SCONJ -> 0.667  ADJ -> 0.654  ADP -> 0.949  ADV -> 0.615  VERB -> 0.826  DET -> 0.838  NOUN -> 0.743  PRON -> 0.806  PROPN -> 0.667  NUM -> 0.957  PUNCT -> 1.000  AUX -> 0.915 |
| 4. | Recall (per tag) | PART -> 1.000  CCONJ -> 1.000  SCONJ -> 1.000  ADJ -> 1.000  ADP -> 0.999  ADV -> 0.982  VERB -> 0.980  DET -> 0.996  NOUN -> 0.998  PRON -> 0.995  PROPN -> 0.996  NUM -> 1.000  PUNCT -> 1.000  AUX -> 1.000 | PART -> 0.909  CCONJ -> 1.000  SCONJ -> 0.667  ADJ -> 0.745  ADP -> 0.971  ADV -> 0.381  VERB -> 0.768  DET -> 0.861  NOUN -> 0.888  PRON -> 0.831  PROPN -> 0.441  NUM -> 0.880  PUNCT -> 0.830  AUX -> 0.935 |
| 5. | F-Score (per tag) | PART -> 1.000  CCONJ -> 0.997  SCONJ -> 0.992  ADJ -> 0.999  ADP -> 0.998  ADV -> 0.982  VERB -> 0.989  DET -> 0.996  NOUN -> 0.998  PRON -> 0.997  PROPN -> 0.998  NUM -> 1.000  PUNCT -> 1.000  AUX -> 0.992 | PART -> 0.909  CCONJ -> 1.000  SCONJ -> 0.667  ADJ -> 0.697  ADP -> 0.960  ADV -> 0.471  VERB -> 0.796  DET -> 0.849  NOUN -> 0.809  PRON -> 0.818  PROPN -> 0.531  NUM -> 0.917  PUNCT -> 0.907  AUX -> 0.925 |
| 6. | Overall Accuracy | 0.9951573849878934 | 0.8334475668265936 |