REPORT - ASSIGNMENT 2

- ∉ Feature Description : In the Assignment, For feature extraction, I used word shape, word suffix, word identity and also some information about the nearby words as features.
- ∉ Hyper parameters tuned (if any) :
  $c1 = 0.005$, $c2 = 0.0267$ for max 150 iterations.

#		Train	Test
1.	10 most common transition features	VERB -> AUX 4.465070 PROPN -> PROPN 2.738170 AUX -> AUX 2.661140 PRON -> ADP 2.628760 ADJ -> NOUN 2.579719 PROPN -> ADP 2.523272A UX -> SCONJ 2.225268 NOUN -> ADP 2.135462 NUM -> NOUN 2.082605 DET -> NOUN 1.979730	VERB -> AUX 4.465070 PROPN -> PROPN 2.738170 AUX -> AUX 2.661140 PRON -> ADP 2.628760 ADJ -> NOUN 2.579719 PROPN -> ADP 2.523272 AUX -> SCONJ 2.225268 NOUN -> ADP 2.135462 NUM -> NOUN 2.082605 DET -> NOUN 1.979730
2.	10 least common transition features	AUX -> ADJ -1.552327 PROPN -> PART -1.603829 DET -> CCONJ -1.641650 PROPN -> DET -1.656438 NUM -> PRON -1.766365 PROPN -> AUX -1.950686 CCONJ -> AUX -1.973740 ADJ -> ADP -2.226438 ADJ -> PRON -2.299370 DET -> ADP -2.824333	AUX -> ADJ -1.552327 PROPN -> PART -1.603829 DET -> CCONJ -1.641650 PROPN -> DET -1.656438 NUM -> PRON -1.766365 PROPN -> AUX -1.950686 CCONJ -> AUX -1.973740 ADJ -> ADP -2.226438 ADJ -> PRON -2.299370 DET -> AD -2.824333
3.	Overall Accuracy	0.99934	0.84072

Training Set -

POS	precision	recall	f1-score	support
X	1	1	1	2
PART	1	1	1	163
CCONJ	1	1	1	150
SCONJ	1	1	1	61
ADJ	1	1	1	569
ADP	0.999	1	1	1384
ADV	1	0.991	0.995	110
VERB	1	0.994	0.997	639
DET	1	1	1	230
COMMA	1	1	1	114
NOUN	1	1	1	1596
PRON	1	1	1	430
PROPN	1	1	1	707
NUM	1	1	1	152
PUNCT	1	1	1	563
AUX	0.995	1	0.997	728
Weighted avg	0.999	0.999	0.999	7598

Test Set -

POS	precision	recall	f1-score	support
X	0	0	0	0
PART	0.968	0.909	0.937	33
CCONJ	1	1	1	25
SCONJ	0.75	1	0.857	3
ADJ	0.667	0.766	0.713	94
ADP	0.946	0.98	0.963	303
ADV	0.6	0.429	0.5	21
VERB	0.859	0.798	0.827	99
DET	0.795	0.861	0.827	36
COMMA	0	0	0	0
NOUN	0.767	0.864	0.813	324
PRON	0.794	0.831	0.812	65
PROPN	0.648	0.486	0.556	144
NUM	0.92	0.92	0.92	25
PUNCT	1	0.828	0.906	134
AUX	0.935	0.942	0.939	138
Weighted avg	0.841	0.841	0.838	1444