REPORT - ASSIGNMENT 4

- Feature Description :

  - Is the first letter of the word capitalised?
  - Is it the first word of the sentence?
  - Is it the last word of the sentence?
  - Is the complete word capitalised?
  - What is the previous and the next word of the word?
  - Is it a number?
  - What are the first four suffixes and prefixes?
  - Does it have a hyphen?

- Hyperparameters tuned (if any) :

  Two fixed hyperparameters 'c1' and 'c2' are declared initially and
  RandomizedSearchCV is used to find the value of the hyperparameters that
  corresponds to the best cross-validation score and those two values of the
  hyperparameters are provided below ,

  `'c1': 0.005378247046498694, 'c2': 0.16508957865144072`

#		Train	Test
1.	10 most common transition features	PROPN -> PROPN 3.885384 VERB -> AUX 3.823892 ADJ -> NOUN 3.489311 NUM -> NOUN 2.380593 DET -> NOUN 2.276857 NOUN -> ADP 2.243182 PROPN -> ADP 2.210423 NOUN -> VERB 1.894061 AUX -> AUX 1.769882 ADJ -> VERB 1.602079	VERB -> AUX 3.080618 PROPN -> PROPN 2.564716 ADJ -> NOUN 2.537360 PROPN -> ADP 2.130822 AUX -> AUX 1.960850 NOUN -> ADP 1.850322 DET -> NOUN 1.713826 NUM -> NOUN 1.663300 NOUN -> VERB 1.465476 NUM -> NUM 1.209100
2.	10 least common transition features	PRON -> CCONJ -1.346362 ADP -> AUX -1.439019 DET -> PROPN -1.452341 ADP -> COMMA -1.478561 DET -> CCONJ -1.480769 ADJ -> PRON -1.594316 CCONJ -> AUX -1.762735 CCONJ -> PART -1.981091 ADJ -> ADP -2.042011 DET -> ADP -2.736387	AUX -> ADP -0.793504 DET -> ADP -0.834493 ADP -> CCONJ -0.854759 VERB -> PROPN -0.959755 PROPN -> ADJ -1.034799 DET -> PROPN -1.110446 AUX -> VERB -1.250181 PROPN -> AUX -1.327473 ADJ -> ADP -1.352714 ADP -> PUNCT -1.474916

3.	Precision (per tag)				
		X	1.000	X	0.000
		PART	1.000	PART	1.000
		CCONJ	0.993	CCONJ	1.000
		SCONJ	0.984	SCONJ	0.750
		ADJ	1.000	ADJ	0.683
		ADP	0.998	ADP	0.961
		ADV	0.982	ADV	0.667
		VERB	1.000	VERB	0.869
		DET	0.996	DET	0.868
		COMMA	1.000	COMMA	0.000
		NOUN	0.999	NOUN	0.795
		PRON	0.995	PRON	0.932
		PROPN	1.000	PROPN	0.673
		NUM	1.000	NUM	0.667
		PUNCT	1.000	PUNCT	1.000
		AUX	0.981	AUX	0.903
4.	Recall (per tag)				
		X	1.000	X	0.000
		PART	1.000	PART	0.970
		CCONJ	1.000	CCONJ	1.000
		SCONJ	1.000	SCONJ	1.000
		ADJ	1.000	ADJ	0.734
		ADP	0.999	ADP	0.977
		ADV	0.982	ADV	0.476
		VERB	0.977	VERB	0.869
		DET	0.991	DET	0.917
		COMMA	1.000	COMMA	0.000
		NOUN	0.999	NOUN	0.886
		PRON	0.995	PRON	0.846
		PROPN	0.996	PROPN	0.514
		NUM	1.000	NUM	0.880
		PUNCT	1.000	PUNCT	0.851
		AUX	1.000	AUX	0.949
5.	F-Score (per tag)				
		X	1.000	X	0.000
		PART	1.000	PART	0.985
		CCONJ	0.997	CCONJ	1.000
		SCONJ	0.992	SCONJ	0.857
		ADJ	1.000	ADJ	0.708
		ADP	0.999	ADP	0.969
		ADV	0.982	ADV	0.556
		VERB	0.988	VERB	0.869
		DET	0.993	DET	0.892
		COMMA	1.000	COMMA	0.000
		NOUN	0.999	NOUN	0.838
		PRON	0.995	PRON	0.887
		PROPN	0.998	PROPN	0.583
		NUM	1.000	NUM	0.759
		PUNCT	1.000	PUNCT	0.919
		AUX	0.990	AUX	0.926
6.	Overall Accuracy	0.9981571673028827		0.8552631578947368	