

# Report on CSE 628 Assignment 1:

## Part 1 :

Evaluation :

### 1. Overall Accuracy:

Overall Accuracy for Part 1: 76.0590195145

### 2. Individual Tags Accuracy:

(u',', 100.0)  
(u'.', 100.0)  
(u'POS', 100.0)  
(u'``, 100.0)  
(u'EX', 100.0)  
(u':', 100.0)  
(u'\$', 100.0)  
(u'WP\$', 100.0)  
(u'-LRB-', 100.0)  
(u'-RRB-', 100.0)  
(u'CC', 99.66101694915255)  
(u'DT', 99.32497589199615)  
(u''''', 99.20634920634922)  
(u'TO', 98.06949806949807)  
(u'IN', 97.82958199356914)  
(u'PRP', 97.52650176678446)  
(u'MD', 96.55172413793103)  
(u'NN', 95.1219512195122)  
(u'WP', 94.11764705882352)  
(u'WRB', 93.54838709677419)  
(u'-NONE-', 90.09661835748793)  
(u'RP', 78.37837837837837)  
(u'PRP\$', 67.76859504132231)  
(u'RB', 67.6829268292683)  
(u'VBP', 67.05882352941175)  
(u'VBZ', 66.24605678233438)  
(u'CD', 64.28571428571429)  
(u'JJ', 62.5)  
(u'VBD', 62.149532710280376)  
(u'RBR', 55.55555555555556)  
(u'JJR', 43.47826086956522)  
(u'VB', 41.69381107491857)  
(u'JJ', 41.340075853350186)

(u'WDT', 41.17647058823529)  
(u'NNP', 40.34511092851273)  
(u'VBG', 36.81318681318682)  
(u'VBN', 34.177215189873415)  
(u'NNS', 33.5149863760218)  
(u'NNPS', 2.7027027027027026)  
(u'PDT', 0.0)  
(u'RBS', 0.0)  
(u'FW', 0.0)

### 3. Confusion matrix :

Confusion matrix for Part 1

θ	JJ	NN	NNP	NNPS	RB	RP	IN	VB	VBD	VBN	VBP
JJ	327	432	2	0	13	0	2	1	0	8	2
NN	13	1521	4	0	0	0	0	43	0	1	7
NNP	7	718	491	1	0	0	0	0	0	0	0
NNPS	0	32	4	1	0	0	0	0	0	0	0
RB	7	72	0	0	222	5	13	1	0	0	0
RP	0	1	0	0	3	29	4	0	0	0	0
IN	2	16	0	0	2	3	1217	0	0	0	0
VB	3	135	0	0	0	0	0	128	0	3	38
VBD	0	115	0	0	0	0	0	1	266	46	0
VBN	0	131	0	0	0	0	0	1	24	81	0
VBP	0	41	0	0	0	0	1	13	0	1	114

#### Analysis:

1. A few tags in the top (u',', 100.0), (u'!', 100.0), (u'POS', 100.0), (u'``', 100.0), (u'EX', 100.0), (u':', 100.0), (u'\$', 100.0), (u'WP\$', 100.0) , (u'-LRB-', 100.0), (u'-RRB-', 100.0), give 100% accuracy. This is because there are just 1 or 2 instances of these tags and they are recognized exactly the same. These are special characters and hence wont be tagged as any other POS. The worst accuracy (u'PDT', 0.0), (u'RBS', 0.0), (u'FW', 0.0) is for those tags which are present in training sentences but are no where to be found in the test sentences.
2. To test this we can count the no of these tags present in the training and the test data and verify the above explanations.

#### Part 2:

##### Method A:

##### Evaluation:

1. Overall Accuracy:

Overall Accuracy for Part 2 Method A: 87.704267809

2. Individual Tags accuracy:

('SNN', 97.89526434477575)

('MISC', 96.30141721396475)

('SRB', 66.28571428571428)

('SVB', 60.51188299817185)

('SJJ', 42.00477326968974)

3. Confusion matrix:

Confusion matrix for Part 2 Method A					
Ø	SNN	SJJ	SRB	SVB	MISC
SNN	3907	20	0	58	6
SJJ	448	352	21	12	5
SRB	73	22	232	1	22
SVB	627	3	0	993	18
MISC	206	2	6	0	5572

Method B:

Evaluation:

1. Overall Accuracy:

Overall Accuracy for Part 2 Method B: 87.7872306591

2. Individual Tags accuracy:

('SNN', 98.3676544450025)

('MISC', 96.2584444829378)

('SRB', 67.0487106017192)

('SVB', 60.07326007326007)

('SJJ', 41.796407185628745)

3. Confusion matrix:

Confusion matrix for Part 2 Method B					
Ø	SNN	SJJ	SRB	SVB	MISC
SNN	3917	23	1	35	6
SJJ	448	349	23	10	5
SRB	74	21	234	0	20
SVB	632	4	0	984	18
MISC	206	2	8	0	5557

Analysis:

1. I had expected Method A to perform better than Method B since Method A was derived from more fine trained data rather than Method B.
2. However, from the overall accuracy we can say that there is just a slight difference between Method A and Method B in which Method B performs slightly better than method A. This might be due to the fact that Method A does the coarse tagging just by matching the tags from the fine predictions without training it on any data whereas Method B first trains the data on the coarse grained tags.