

COMS 4705 - Homework 1

Aayush Maini - am4810

Programming Problems

Question 4

	Precision	Recall	F1-Score
Total	0.22	0.52	0.31
PER	0.44	0.23	0.30
ORG	0.48	0.40	0.43
LOC	0.15	0.87	0.25
MISC	0.49	0.61	0.54

Table 1: Performance of 4.2.py

This baseline tagger performs very poorly because it doesn't take in account any context. Further, the number of NEs found are around 14,000. This is due to the reason that since the baseline tagger doesn't take into the the context, it assigns incompatible tags to adjacent words. To explain, these incompatible tags, new Named Entities will have to be created.

Question 5

	Precision	Recall	F1-Score
Total	0.77	0.61	0.68
PER	0.76	0.59	0.67
ORG	0.61	0.48	0.54
LOC	0.88	0.70	0.78
MISC	0.83	0.69	0.75

Table 2: Performance of 5.2.py

The Viterbi tagger takes into account the context information and hence, performs better than the baseline tagger in 4.2.py

Question 6

	Precision	Recall	F1-Score
Total	0.74	0.73	0.74
PER	0.81	0.78	0.79
ORG	0.54	0.66	0.60
LOC	0.84	0.75	0.80
MISC	0.83	0.68	0.75

Table 3: Performance of 6.py

Class	Example
_TWO_DIGIT_NUM_	12
_FOUR_DIGIT_NUM_	1234
_DIGIT_AND_ALPHA_	ABC123
_DIGIT_AND_DASH_	Date in US and European formats
_DIGIT_AND_PERIOD_	123.34
_OTHER_NUM_	Any other number
_ALL_CAPS_	ABCDE
_CAP_PERIOD_	P. (usually initials of first name)
_INIT_CAP_	Capitalized word like "This"
LOWERCASE	Any lowercase word
OTHER	Any other rare word

Table 4: Rare word classification

- Excluding the Date category, improves performance by a very small margin. This might be because the tagger is overfitting to the current training data. Similarly, excluding the _TWO_DIGIT_NUM_ and _FOUR_DIGIT_NUM_ category improves the performance

	Precision	Recall	F1-Score
Total	0.74	0.73	0.73
PER	0.80	0.77	0.79
ORG	0.54	0.66	0.59
LOC	0.84	0.75	0.79
MISC	0.82	0.67	0.74

- Excluding the `_ALL_CAPS_`, `_CAP_PERIOD_`, `_INIT_CAP_` and `_LOWERCASE_` category decreases the performance by a huge amount. This is because these words are significant in number in the training data and hence, they do render the algorithm susceptible to high variance and overfitting

	Precision	Recall	F1-Score
Total	0.77	0.62	0.68
PER	0.74	0.60	0.66
ORG	0.63	0.50	0.56
LOC	0.87	0.69	0.77
MISC	0.82	0.68	0.75