

COL776 (Assignment -2)

J. Shikhar Murty

Implementation Language: Python

Experiments:

1) Comparing Inference Algorithms:

data-loops.dat

Algorithm	Char accuracy	Word accuracy	Log likelihood	Time taken
MP	56.8345	7.1429	-7.5454	0.559
LBP	56.8345	7.1429	-7.5454	7.311

data-loopsWS.dat

Algorithm	Char accuracy	Word accuracy	Log likelihood	Time taken
MP	67.5926	13.8462	-6.9198	2.161
LBP	67.9012	15.3846	-6.9095	34.21

data-tree.dat

Algorithm	Char accuracy	Word accuracy	Log likelihood	Time taken
MP	67.5603	16.6667	-6.6065	0.42
LBP	67.5603	16.6667	-6.6065	21.80

data-treeWS.dat

Algorithm	Char accuracy	Word accuracy	Log likelihood	Time taken
MP	66.6667	16.3043	-7.2087	2.01
LBP	66.7770	16.8478	-7.2095	48.10

Comments

- 1) LBP and MP have very similar performance for all datasets.
- 2) Instead of running LBP till convergence, the stopping criteria used was to run for a fixed time ($\frac{1}{2}$ second).
- 3) LBP takes more time than MP for similar performance.

2) Adding sophistication in factors

data-loopsWS.dat

Model	Char accuracy	Word accuracy	Log likelihood	Time taken
OCR	60.0309	8.4615	-7.8162	0.1845
OCR+T	66.0494	11.5385	-7.5249	0.6117
OCR+T+S	66.2037	11.5385	-7.3318	1.3255
OCR+T+S+PS	67.5926	13.8462	-6.9198	2.1511

We observe that on adding sophistication, the performance increases, as expected. Also note that the time taken increases since the tree width of the underlying MRF also increases. However, the increase in performance is not as evident as compared to the MAP inference case, as we'll soon see.

MAP Inference:

1) Comparing Inference Algorithms

data-loops.dat

Algorithm	Char accuracy	Word accuracy	Log likelihood	Time taken
MP	55.3957	17.8571	-7.0556	0.537
LBP	56.8345	17.8571	-7.1025	7.326

data-loopsWS.dat

Algorithm	Char accuracy	Word accuracy	Log likelihood	Time taken
MP	80.0926	62.3077	-6.2662	2.162
LBP	79.1667	59.2308	-6.0826	34.14

data-tree.dat

Algorithm	Char accuracy	Word accuracy	Log likelihood	Time taken
MP	77.2118	52.3810	-6.1714	0.408
LBP	77.2118	52.3803	-6.1714	21.72

data-treeWS.dat

Algorithm	Char accuracy	Word accuracy	Log likelihood	Time taken
MP	76.4901	50.0000	-6.4832	2.054
LBP	76.6004	48.3696	-6.5071	48.27

2) Adding sophistication in factors

data-loopsWS.dat

Model	Char accuracy	Word accuracy	Log likelihood	Time taken
OCR	60.0309	8.4615	-7.8162	0.1915
OCR+T	74.3827	34.6154	-7.3170	0.5868
OCR+T+S	76.8519	50.7692	-6.7895	1.6118
OCR+T+S+PS	80.0926	62.3077	-6.2662	2.5675

Comments:

1)Adding Skip Factors significantly increases the word accuracy, and adding pair skip factors further increases the word accuracy, as expected.

2)Similar to max marginal inference, the time taken increases with increased model complexity as the tree width of the corresponding markov network also increases.