NLP

POS tagging using HMM & MLP models

Contents

HMM

Fold number	Accurary for folds	Precision	Recall	F1 score	Total words	Tags found
1	89.9	0.86	0.84	0.84	6946	12
2	90.46	0.91	0.88	0.89	7018	12
3	90.46	0.90	0.88	0.89	6628	12
Overall Accuracy	89.86	0.89	0.89	0.87	0.88	12

MLP

using word2vec

	Accuracy	Precision	Recall	F1 score
	88.52	0.82	0.81	0.81

Using Glove

Fold number	Accuracy	Precision	Recall	F1 score
1	87.94	0.89	0.84	0.86
2	87.96	0.92	0.86	0.83
3	0.91	0.81	0.81	0.84

Comparative study of MLP and HMM

Task	HMM	MLP (Glove)

Precision	0.86	0.86
Recall	0.84	0.82
F measure	0.84	0.82
Accuracy	0.91	0.90

Tag wise precision /recall and F1 score using HMM and MLP

Task	HMM	MLP
1	0.86	0.91
2	0.82	0.81
3	0.82	0.84

Statistics of tag set using HMM:

ADP 7327
DET 6024
NOUN 14268
VERB 6968
ADJ 4135
CONJ 1701
PRT 883
. 5170
ADV 1620
NUM 1194
PRON 553
X 20

Incorrect classifications

No of incorrect classifications are 2305

Reasons for incorrect classifications is Ambiguity,The main problem with POS tagging is **ambiguity**. In English, many common words have multiple meanings and therefore multiple POS . The job of a POS tagger is to resolve this ambiguity accurately based on the context of use. For example, the word "shot" can be a noun or a verb

{'tags': [[4, 'NOUN', 'ADP'], [5, 'VERB', 'DET']], 'text': ['At', 'that', 'time', 'highway', 'engineers', 'traveled', 'rough', 'and', 'dirty', 'roads', 'to', 'accomplish', 'their', 'duties', '.']}

{'tags': [[0, 'VERB', 'DET'], [20, 'ADV', 'DET'], [21, 'ADJ', 'NOUN']], 'text': ['Using', 'privately-owned', 'vehicles', 'was', 'a', 'personal', 'hardship', 'for', 'such', 'employees', ',', 'and', 'the', 'matter', 'of', 'providing', 'state', 'transportation', 'was', 'felt', 'perfectly', 'justifiable', '.']}

{'tags': [[9, 'ADJ', 'DET']], 'text': ['Once', 'the', 'principle', 'was', 'established', ',', 'the', 'increase', 'in', 'state-owned', 'vehicles', 'came', 'rapidly', '.']}

{'tags': [[6, 'VERB', 'ADJ']], 'text': ['This', 'rate', 'of', 'increase', 'does', 'not', 'signify', 'anything', 'in', 'itself', '.']}

{'tags': [[4, 'ADJ', 'DET'], [7, 'ADJ', 'DET']], 'text': ['It', 'does', 'not', 'indicate', 'loose', 'management', ',', 'ineffective', 'controls', 'or', 'poor', 'policy', '.']}

{'tags': [[5, 'NOUN', 'DET']], 'text': ['But', 'it', 'does', 'show', 'that', 'automobiles', 'have', 'increased', 'steadily', 'over', 'the', 'years', 'and', 'in', 'almost', 'the', 'same', 'proportion', 'to', 'the', 'increase', 'of', 'state', 'employees', '.']}

{'tags': [[2, 'ADJ', 'NOUN'], [8, 'ADJ', 'DET']], 'text': ['In', 'the', 'past', 'twenty', 'years', 'the', 'ratio', 'of', 'state-owned', 'automobiles', 'per', 'state', 'employees', 'has', 'varied', 'from', '1', 'to', '22', 'then', 'to', '1', 'to', '23', 'now', '.']}

{'tags': [[7, 'NUM', 'NOUN'], [13, 'ADJ', 'NOUN']], 'text': ['Whether', 'there', 'were', 'too', 'few', 'automobiles', 'in', '1940', 'or', 'too', 'many', 'now', 'is', 'problematical', '.']}

Task 1 - HMM

Overall Statistics

Final Accuracy = 89.86805098147188
Final Precision = 0.8944755633450088
Final Recall = 0.8729588603886583
Final F1 Score = 0.8806971415645234
Total sentencess = 27491
Maximum length sentences = 386
Minimum length sentences = 2
Average length of sentences= 19.756502127969153

Task 2 - MLP

Word2vec - Overall Statistics

Final Accuracy = 88.14908257601871
Final Precision = 0.8643173565813945
Final Recall = 0.8207729458549986
Final F1 Score = 0.8271486451739071
Total sentencess = 27491
Maximum length sentences = 386
Minimum length sentences = 2
Average length of sentences= 19.756502127969153

GLoVE - Overall Statistics

Final Accuracy = 87.74926774288456
Final Precision = 0.9151915057494775
Final Recall = 0.8197749145175212
Final F1 Score = 0.8457407565144185
Total sentences = 27491
Maximum length sentences = 386
Minimum length sentences = 2
Average length of sentences= 19.7565

MLP is better than HMM as the final accuracy of MLP,precision and recall across all the folds of the data gives better results. Even the MLP provides better precision for tag set and also incorrect tag descriptions are limited.

Other statistics

Words having more than 1 tag
that {'DET': 1095, 'ADP': 2789, 'PRON': 670, 'ADV': 33}
time {'NOUN': 776, 'VERB': 1}
to {'PRT': 6763, 'ADP': 5089, 'X': 1}
a {'DET': 9819, 'X': 2, 'NOUN': 1}
for {'ADP': 3898, 'ADV': 2}
such {'ADJ': 431, 'PRT': 142, 'ADV': 9}
matter {'NOUN': 117, 'VERB': 20}
of {'ADP': 15791, 'X': 2}
state {'NOUN': 243, 'VERB': 6}
Once {'ADP': 14, 'ADV': 36}
principle {'NOUN': 43, 'ADJ': 1}
increase {'NOUN': 76, 'VERB': 37}
in {'ADP': 8321, 'PRT': 278, 'X': 1}
reasons {'NOUN': 38, 'VERB': 1}
need {'NOUN': 64, 'VERB': 66}
1 {'NUM': 344, 'NOUN': 1}
below {'ADV': 43, 'ADP': 45}
shows {'VERB': 38, 'NOUN': 1}
loose {'ADJ': 19, 'VERB': 1, 'ADV': 2}
But {'CONJ': 610, 'ADP': 1}
it {'PRON': 3376, 'PRT': 1}
show {'VERB': 101, 'NOUN': 17}
over {'ADP': 419, 'PRT': 207, 'ADJ': 3}
In {'ADP': 813, 'PRT': 1}
past {'ADJ': 45, 'NOUN': 36, 'ADP': 38, 'ADV': 8}
then {'ADV': 562, 'ADJ': 10}
there {'PRT': 656, 'ADV': 341}
many {'ADJ': 317, 'ADV': 4, 'PRT': 7}
as {'ADP': 2505, 'ADV': 444}
use {'VERB': 88, 'NOUN': 166}
(this is a long list, which can be generated through code)
……