

By adding those features I've seen improvements in both dev and dev_test, first I added all of files and tried to check every word in those files and list it seemed that those features are too much and kind of overfited when it comes to dev test, then as I explained in 1.1 part I removed some words and did some reading on each file and tried to just extract tags from words that are related. Feature like adding just first letter of the word seemed didn't work because of overfitting so I removed that too. Then after that I added checking two words when I'm checking the sentence, it means I check not only a word after and before the current word but also words that are around them, then my features became around 74k but my accuracy got higher in both dev and dev_test by 3 and 2 percent.

After adding features:

```
Anaconda Prompt - python main.py -e data/twitter_dev.ner data/twitter_dev_test.ner
Features computed
(34932, 73404)
### Train evaluation; writing to ./twitter_train.ner.pred
Token-wise accuracy 99.086797206
Token-wise F1 (macro) 89.0264318729
Token-wise F1 (micro) 99.086797206
Sentence-wise accuracy 88.5809312639
```

	precision	recall	f1-score	support
B-company	1.00	0.81	0.90	135
B-facility	1.00	0.86	0.92	76
B-geo-loc	1.00	0.87	0.93	199
B-movie	1.00	0.81	0.90	27
B-musicartist	1.00	0.76	0.86	42
B-other	0.99	0.77	0.87	162
B-person	1.00	0.89	0.94	341
B-product	1.00	0.73	0.84	78
B-sportsteam	1.00	0.72	0.84	40
B-tvshow	1.00	0.78	0.88	23
I-company	1.00	0.86	0.93	29
I-facility	1.00	0.88	0.94	76
I-geo-loc	1.00	0.80	0.89	35
I-movie	1.00	0.66	0.79	35
I-musicartist	1.00	0.80	0.89	46
I-other	0.98	0.82	0.89	239
I-person	1.00	0.92	0.96	154
I-product	1.00	0.73	0.85	64
I-sportsteam	1.00	0.79	0.88	19
I-tvshow	1.00	0.67	0.80	21
o	0.99	1.00	1.00	33091
avg / total	0.99	0.99	0.99	34932

```

### evaluation of data/twitter_dev.ner; writing to ./twitter_dev.ner.pred
Token-wise accuracy 95.8134697062
Token-wise F1 (macro) 26.3301139566
Token-wise F1 (micro) 95.8134697062
Sentence-wise accuracy 67.4576271186

```

	precision	recall	f1-score	support
B-company	0.85	0.31	0.45	36
B-facility	0.69	0.39	0.50	28
B-geo-loc	0.76	0.34	0.47	77
B-movie	0.00	0.00	0.00	7
B-musicartist	0.00	0.00	0.00	13
B-other	0.67	0.13	0.21	63
B-person	0.68	0.42	0.52	108
B-product	1.00	0.11	0.19	19
B-sportsteam	0.00	0.00	0.00	11
B-tvshow	0.50	0.09	0.15	11
I-company	0.00	0.00	0.00	7
I-facility	0.58	0.38	0.46	29
I-geo-loc	0.33	0.07	0.12	14
I-movie	0.00	0.00	0.00	11
I-musicartist	1.00	0.07	0.12	15
I-other	0.64	0.20	0.30	81
I-person	0.75	0.34	0.47	61
I-product	1.00	0.25	0.40	16
I-sportsteam	0.00	0.00	0.00	4
I-tvshow	1.00	0.10	0.18	10
O	0.96	1.00	0.98	10916
avg / total	0.95	0.96	0.95	11537

```

### evaluation of data/twitter_dev_test.ner; writing to ./twitter_dev_test.ner.pred
Token-wise accuracy 91.5546515741
Token-wise F1 (macro) 17.8367958532
Token-wise F1 (micro) 91.5546515741
Sentence-wise accuracy 49.5021337127

```

	precision	recall	f1-score	support
B-company	0.75	0.08	0.15	109
B-facility	0.58	0.24	0.34	46
B-geo-loc	0.72	0.35	0.47	159
B-movie	0.00	0.00	0.00	4
B-musicartist	0.00	0.00	0.00	33
B-other	0.25	0.03	0.05	118
B-person	0.37	0.31	0.34	96
B-product	0.00	0.00	0.00	44
B-sportsteam	0.00	0.00	0.00	31
B-tvshow	0.00	0.00	0.00	4
I-company	1.00	0.04	0.07	26
I-facility	0.67	0.27	0.38	60
I-geo-loc	0.82	0.24	0.38	37
I-movie	0.00	0.00	0.00	10
I-musicartist	0.00	0.00	0.00	15
I-other	0.44	0.17	0.25	123
I-person	0.49	0.29	0.37	58
I-product	0.00	0.00	0.00	88
I-sportsteam	0.00	0.00	0.00	7
I-tvshow	0.00	0.00	0.00	9
O	0.93	1.00	0.96	10231
avg / total	0.88	0.92	0.89	11308

Before adding features:

```
Features computed
(34932, 45376)
### Train evaluation; writing to ./twitter_train.ner.pred
Token-wise accuracy 97.7728157563
Token-wise F1 (macro) 66.9137520549
Token-wise F1 (micro) 97.7728157563
Sentence-wise accuracy 76.7738359202
      precision    recall  f1-score   support

   B-company      0.99      0.60      0.75      135
  B-facility      0.96      0.63      0.76       76
  B-geo-loc       0.94      0.66      0.78      199
    B-movie       1.00      0.52      0.68       27
B-musicartist     1.00      0.40      0.58       42
    B-other       0.96      0.51      0.66      162
    B-person      0.98      0.70      0.82      341
    B-product     1.00      0.41      0.58       78
  B-sportsteam    1.00      0.35      0.52       40
    B-tvshow      1.00      0.30      0.47       23
    I-company     1.00      0.59      0.74       29
  I-facility      1.00      0.51      0.68       76
  I-geo-loc       1.00      0.54      0.70       35
    I-movie       1.00      0.40      0.57       35
I-musicartist     1.00      0.37      0.54       46
    I-other       0.96      0.58      0.72      239
    I-person      0.99      0.79      0.88      154
    I-product     1.00      0.38      0.55       64
  I-sportsteam    1.00      0.42      0.59       19
    I-tvshow      1.00      0.33      0.50       21
         0        0.98      1.00      0.99    33091

 avg / total      0.98      0.98      0.97    34932
```

```
### evaluation of data/twitter_dev_test.ner; writing to ./twitter_dev_test
```

```
Token-wise accuracy 91.0152104705
```

```
Token-wise F1 (macro) 10.9195384447
```

```
Token-wise F1 (micro) 91.0152104705
```

```
Sentence-wise accuracy 48.6486486486
```

	precision	recall	f1-score	support
B-company	0.82	0.08	0.15	109
B-facility	0.56	0.30	0.39	46
B-geo-loc	0.72	0.30	0.42	159
B-movie	0.00	0.00	0.00	4
B-musicartist	0.00	0.00	0.00	33
B-other	0.00	0.00	0.00	118
B-person	0.22	0.10	0.14	96
B-product	0.00	0.00	0.00	44
B-sportsteam	0.00	0.00	0.00	31
B-tvshow	0.00	0.00	0.00	4
I-company	0.00	0.00	0.00	26
I-facility	0.00	0.00	0.00	60
I-geo-loc	1.00	0.03	0.05	37
I-movie	0.00	0.00	0.00	10
I-musicartist	0.00	0.00	0.00	15
I-other	0.62	0.08	0.14	123
I-person	0.14	0.02	0.03	58
I-product	0.00	0.00	0.00	88
I-sportsteam	0.00	0.00	0.00	7
I-tvshow	0.00	0.00	0.00	9
0	0.92	1.00	0.95	10231
avg / total	0.86	0.91	0.88	11308

evaluation of data/twitter_dev.ner; writing to ./twitter_dev.ner.pred

Token-wise accuracy 95.5361012395

Token-wise F1 (macro) 21.5780375334

Token-wise F1 (micro) 95.5361012395

Sentence-wise accuracy 66.6101694915

	precision	recall	f1-score	support
B-company	1.00	0.33	0.50	36
B-facility	0.62	0.36	0.45	28
B-geo-loc	0.82	0.30	0.44	77
B-movie	0.00	0.00	0.00	7
B-musicartist	0.00	0.00	0.00	13
B-other	1.00	0.10	0.17	63
B-person	0.70	0.29	0.41	108
B-product	1.00	0.16	0.27	19
B-sportsteam	0.00	0.00	0.00	11
B-tvshow	0.00	0.00	0.00	11
I-company	0.00	0.00	0.00	7
I-facility	0.67	0.14	0.23	29
I-geo-loc	1.00	0.07	0.13	14
I-movie	0.00	0.00	0.00	11
I-musicartist	0.00	0.00	0.00	15
I-other	0.58	0.14	0.22	81
I-person	0.92	0.20	0.32	61
I-product	1.00	0.25	0.40	16
I-sportsteam	0.00	0.00	0.00	4
I-tvshow	0.00	0.00	0.00	10
0	0.96	1.00	0.98	10916
avg / total	0.94	0.96	0.94	11537