

### 3. COMPARE LOGISTIC REGRESSION TO CRFs

#### 3.2 EXPERIEMENTS

##### Conlleval Script Result Evaluation

		BASIC FEATURES		ADVANCED FEATURES	
		LogReg	CRF	LogReg	CRF
Twitter_dev.ner	Accuracy	95.54	95.77	95.86	96.06
	Precision	49.61	60.61	46.68	59.36
	Recall	16.89	26.81	24.66	34.85
	FB1	25.20	37.17	32.74	43.92
Twitter_dev_test.ner	Accuracy	91.02	91.31	91.65	91.91
	Precision	32.35	46.82	35.22	46.08
	Recall	8.54	15.99	17.39	23.76
	FB1	13.51	23.84	23.38	31.35

The Conlleval statistics for the methods Logistic Regression and CRF, and datasets twitter\_dev.ner and twitter\_dev\_test.ner is shown above.

In the Conlleval evaluation, the 'accuracy' can be used as the aggregate measure to compare the performance of Logistic Regression and CRF.

For basic and advanced feature set, CRF performs better than Logistic Regression.

For the basic feature set and for twitter\_dev.ner dataset, the accuracy of CRF exceeds that of Logistic Regression by 0.23. Hence, the accuracy is only slightly higher in CRF than in Logistic Regression. However, with respect to precision, recall and FB1, the CRF performance exceeds Logistic Regression by 11, 9.92 and 11.97 respectively. For twitter\_dev\_test.ner dataset, the accuracy of CRF exceeds that of Logistic Regression by 0.29. Again, the accuracy is only slightly higher in CRF than in Logistic Regression. However, with respect to precision, recall and FB1, the CRF performance exceeds Logistic Regression by 14.47, 7.45 and 10.33 respectively.

For the advanced feature set and twitter\_dev.ner dataset, the accuracy of CRF exceeds that of Logistic Regression by 0.2. However, with respect to precision, recall and FB1, the CRF performance exceeds Logistic Regression by 12.68, 10.19 and 11.18 respectively. For twitter\_dev\_test.ner dataset, the accuracy of CRF exceeds that of Logistic Regression by 0.26. Again, the accuracy is only slightly higher in CRF than in Logistic Regression. However, with respect to precision, recall and FB1, the CRF performance exceeds Logistic Regression by 10.86, 6.37 and 7.97 respectively.

CRF performs better than Logistic Regression as per the above statistics obtained. The primary reason why CRF performs better would be because it does not assume that the features of the model are independent. While assigning a label to a word, it takes into account the predictions

of the surrounding words. Also, the performance of the CRF model is greatly tunable by the features.

The screenshots of CONLL scripts are:

### Logistic Regression – Basic Features

```
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$ perl data/conlleval.pl -d \\t < twitter_dev.ner.pred
processed 11537 tokens with 373 phrases; found: 127 phrases; correct: 63.
accuracy: 95.54%; precision: 49.61%; recall: 16.89%; FB1: 25.20
  company: precision: 100.00%; recall: 33.33%; FB1: 50.00 12
  facility: precision: 10.00%; recall: 7.14%; FB1: 8.33 20
  geo-loc: precision: 78.57%; recall: 28.57%; FB1: 41.90 28
  movie: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
  musicartist: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
  other: precision: 15.79%; recall: 4.76%; FB1: 7.32 19
  person: precision: 46.67%; recall: 19.44%; FB1: 27.45 45
  product: precision: 100.00%; recall: 15.79%; FB1: 27.27 3
  sportsteam: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
  tvshow: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$ perl data/conlleval.pl -d \\t < twitter_dev_test.ner.pred
processed 11308 tokens with 644 phrases; found: 170 phrases; correct: 55.
accuracy: 91.02%; precision: 32.35%; recall: 8.54%; FB1: 13.51
  company: precision: 72.73%; recall: 7.34%; FB1: 13.33 11
  facility: precision: 4.00%; recall: 2.17%; FB1: 2.82 25
  geo-loc: precision: 58.21%; recall: 24.53%; FB1: 34.51 67
  movie: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
  musicartist: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
  other: precision: 0.00%; recall: 0.00%; FB1: 0.00 17
  person: precision: 14.29%; recall: 7.29%; FB1: 9.66 49
  product: precision: 0.00%; recall: 0.00%; FB1: 0.00 1
  sportsteam: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
  tvshow: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$
```

### Logistic Regression – Advanced Features

```
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$ perl data/conlleval.pl -d \\t < twitter_dev.ner.pred
processed 11537 tokens with 373 phrases; found: 189 phrases; correct: 92.
accuracy: 95.86%; precision: 48.68%; recall: 24.66%; FB1: 32.74
  company: precision: 80.00%; recall: 33.33%; FB1: 47.06 15
  facility: precision: 11.54%; recall: 10.71%; FB1: 11.11 26
  geo-loc: precision: 83.33%; recall: 38.96%; FB1: 53.10 36
  movie: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
  musicartist: precision: 0.00%; recall: 0.00%; FB1: 0.00 2
  other: precision: 25.93%; recall: 11.11%; FB1: 15.56 27
  person: precision: 48.65%; recall: 33.33%; FB1: 39.56 74
  product: precision: 60.00%; recall: 15.79%; FB1: 25.00 5
  sportsteam: precision: 0.00%; recall: 0.00%; FB1: 0.00 2
  tvshow: precision: 50.00%; recall: 9.09%; FB1: 15.38 2
Nehas-MacBook-Air:Homework-6 nehapathapati$ perl data/conlleval.pl -d \\t < twitter_dev_test.ner.pred
processed 11308 tokens with 644 phrases; found: 318 phrases; correct: 112.
accuracy: 91.65%; precision: 35.22%; recall: 17.39%; FB1: 23.28
  company: precision: 75.00%; recall: 13.76%; FB1: 23.26 20
  facility: precision: 11.11%; recall: 8.70%; FB1: 9.76 36
  geo-loc: precision: 65.71%; recall: 43.40%; FB1: 52.27 105
  movie: precision: 0.00%; recall: 0.00%; FB1: 0.00 2
  musicartist: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
  other: precision: 0.00%; recall: 0.00%; FB1: 0.00 36
  person: precision: 19.83%; recall: 23.96%; FB1: 21.70 116
  product: precision: 100.00%; recall: 2.27%; FB1: 4.44 1
  sportsteam: precision: 0.00%; recall: 0.00%; FB1: 0.00 2
  tvshow: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
Nehas-MacBook-Air:Homework-6 nehapathapati$
```

## CRF – Basic Features

```
Running setup.py install for perl ... done
Successfully installed perl-0.1
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$ sudo data/ conlleval.pl -d \\t < twitter_dev.ner.pred
[sudo: data/ conlleval.pl: command not found
Nehas-MacBook-Air:Homework-6 nehapathapati$ data/ conlleval.pl -d \\t < twitter_dev.ner.pred
[-bash: data/ conlleval.pl: Permission denied
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$ perl data/ conlleval.pl -d \\t < twitter_dev.ner.pred
processed 11537 tokens with 373 phrases; found: 165 phrases; correct: 100.
accuracy: 95.77%; precision: 60.61%; recall: 26.81%; FB1: 37.17
company: precision: 87.50%; recall: 38.89%; FB1: 53.85 16
facility: precision: 47.62%; recall: 35.71%; FB1: 40.82 21
geo-loc: precision: 70.00%; recall: 36.36%; FB1: 47.86 40
movie: precision: 0.00%; recall: 0.00%; FB1: 0.00 1
musicartist: precision: 0.00%; recall: 0.00%; FB1: 0.00 2
other: precision: 53.85%; recall: 11.11%; FB1: 18.42 13
person: precision: 57.38%; recall: 32.41%; FB1: 41.42 61
product: precision: 60.00%; recall: 15.79%; FB1: 25.00 5
sportsteam: precision: 50.00%; recall: 9.09%; FB1: 15.38 2
tvshow: precision: 50.00%; recall: 18.18%; FB1: 26.67 4
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$ perl data/ conlleval.pl -d \\t < twitter_dev_test.ner.pred
processed 11308 tokens with 644 phrases; found: 220 phrases; correct: 103.
accuracy: 91.31%; precision: 46.82%; recall: 15.99%; FB1: 23.84
company: precision: 82.35%; recall: 12.84%; FB1: 22.22 17
facility: precision: 38.71%; recall: 26.09%; FB1: 31.17 31
geo-loc: precision: 72.84%; recall: 37.11%; FB1: 49.17 81
movie: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
musicartist: precision: 0.00%; recall: 0.00%; FB1: 0.00 1
other: precision: 14.29%; recall: 1.69%; FB1: 3.03 14
person: precision: 21.43%; recall: 15.62%; FB1: 18.07 70
product: precision: 20.00%; recall: 2.27%; FB1: 4.08 5
sportsteam: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
tvshow: precision: 0.00%; recall: 0.00%; FB1: 0.00 1
Nehas-MacBook-Air:Homework-6 nehapathapati$
```

## CRF – Advanced Features

```
main.py~ viterbi_test.py
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$ perl data/ conlleval.pl -d \\t < twitter_dev.ner.pred
processed 11537 tokens with 373 phrases; found: 219 phrases; correct: 130.
accuracy: 96.06%; precision: 59.36%; recall: 34.85%; FB1: 43.92
company: precision: 77.78%; recall: 38.89%; FB1: 51.85 18
facility: precision: 54.55%; recall: 42.86%; FB1: 48.00 22
geo-loc: precision: 70.59%; recall: 46.75%; FB1: 56.25 51
movie: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
musicartist: precision: 33.33%; recall: 7.69%; FB1: 12.50 3
other: precision: 50.00%; recall: 17.46%; FB1: 25.88 22
person: precision: 55.43%; recall: 47.22%; FB1: 51.00 92
product: precision: 100.00%; recall: 15.79%; FB1: 27.27 3
sportsteam: precision: 20.00%; recall: 9.09%; FB1: 12.50 5
tvshow: precision: 33.33%; recall: 9.09%; FB1: 14.29 3
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$
Nehas-MacBook-Air:Homework-6 nehapathapati$ perl data/ conlleval.pl -d \\t < twitter_dev_test.ner.pred
processed 11308 tokens with 644 phrases; found: 332 phrases; correct: 153.
accuracy: 91.91%; precision: 46.08%; recall: 23.76%; FB1: 31.35
company: precision: 66.67%; recall: 12.84%; FB1: 21.54 21
facility: precision: 39.47%; recall: 32.61%; FB1: 35.71 38
geo-loc: precision: 66.96%; recall: 48.43%; FB1: 56.20 115
movie: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
musicartist: precision: 0.00%; recall: 0.00%; FB1: 0.00 0
other: precision: 0.00%; recall: 0.00%; FB1: 0.00 12
person: precision: 36.00%; recall: 46.88%; FB1: 40.72 125
product: precision: 14.29%; recall: 2.27%; FB1: 3.92 7
sportsteam: precision: 8.33%; recall: 3.23%; FB1: 4.65 12
tvshow: precision: 0.00%; recall: 0.00%; FB1: 0.00 2
Nehas-MacBook-Air:Homework-6 nehapathapati$
```