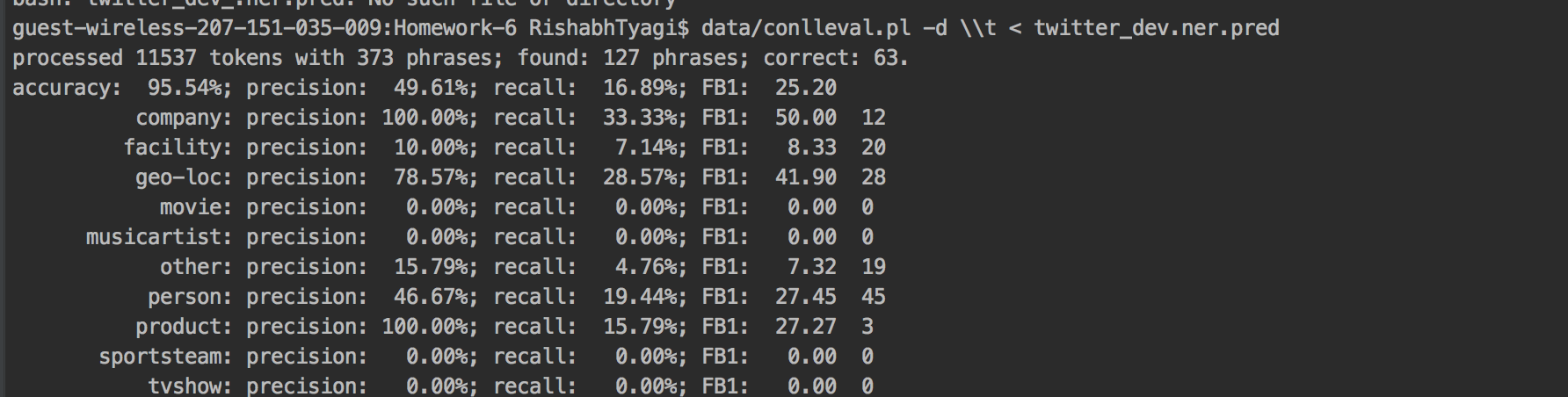
Question 3.2

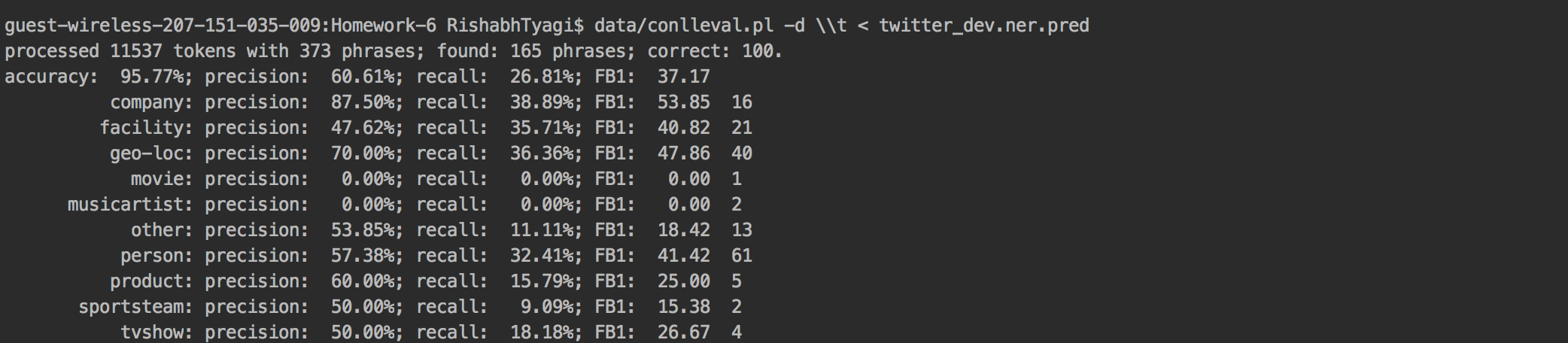
1. Comparing the difference in the output of the evaluation between logistic regression and conditional random field is shown below using just the base features.

**Logistic Regression (without features):** To remove redundancy, I have just added the result from the conlleval evaluation.

**CONLL evaluation for twitter\_dev.ner.pred(without features)**

**CRF: Conditional Random Field (without features):**

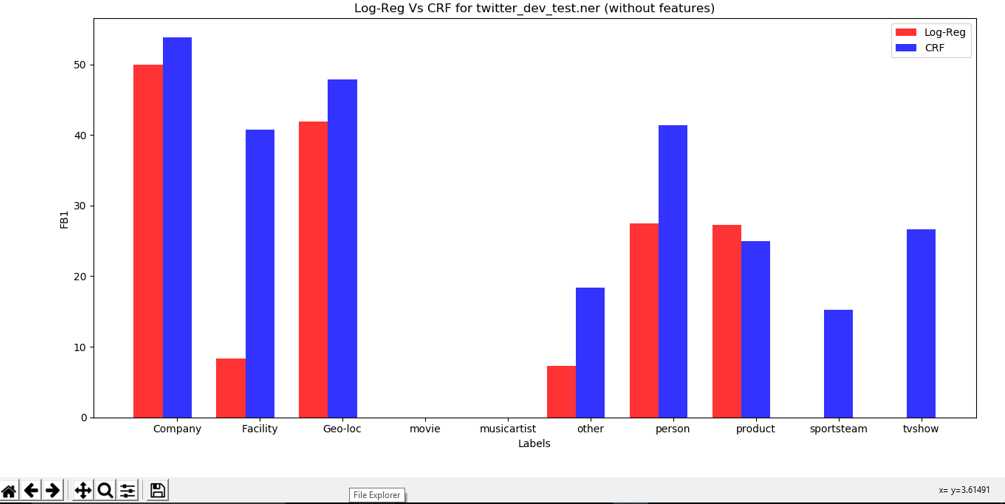
**CONLL evaluation for twitter\_dev.ner.pred(without features)**

****

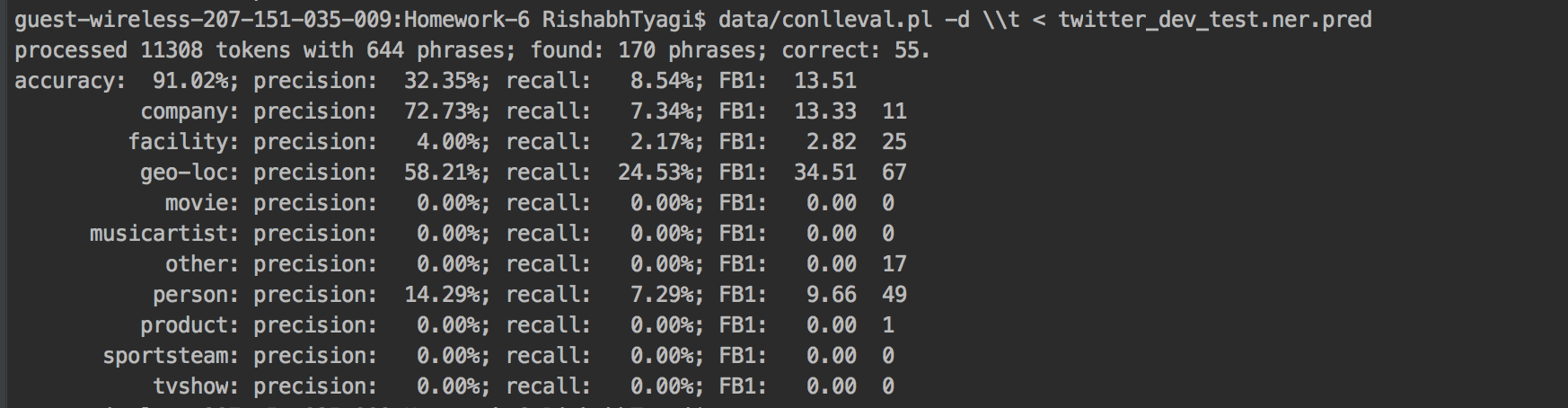
Accuracy of Logistic Regression: 95.54%; FB1: 25.20

Accuracy of Conditional Random Field: 95.77%; FB1: 37.17

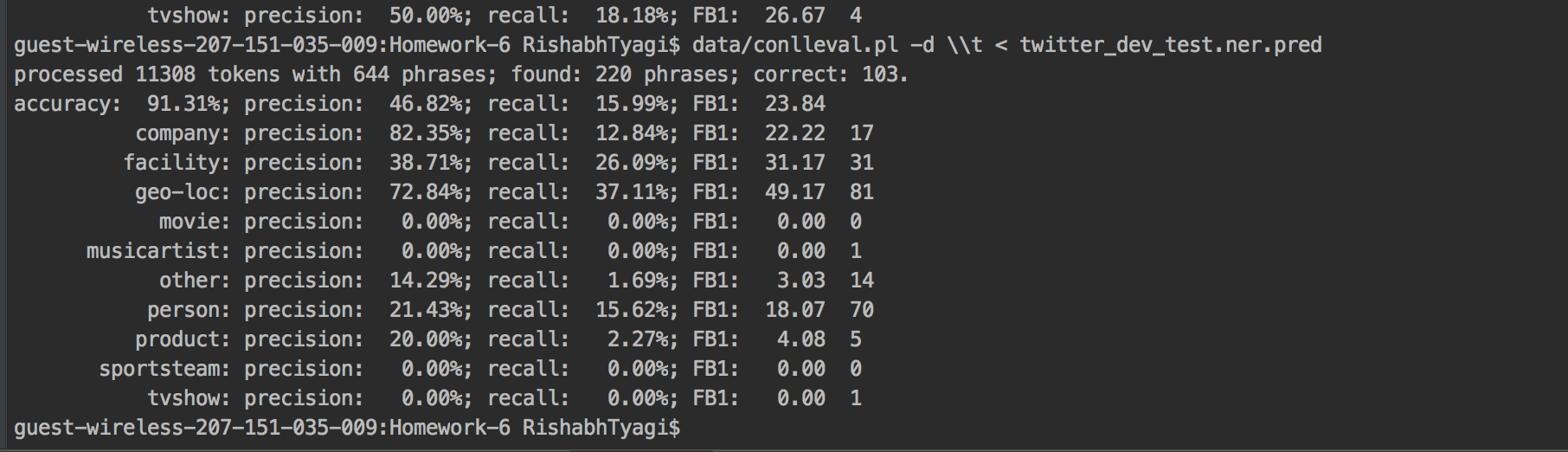
**Graph for CRF vs LogReg for twitter\_dev.ner (without features):**



**CONLL evaluation for twitter\_dev\_test.ner.pred(without features)**



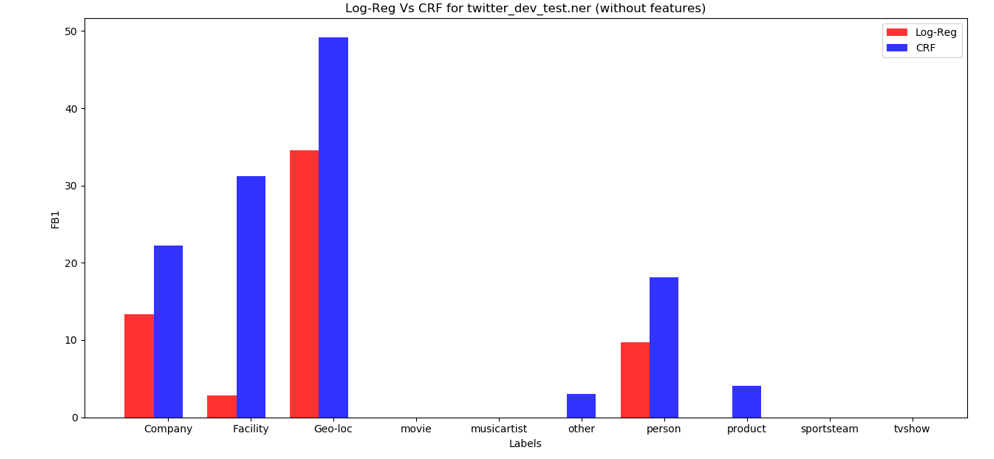
**CONLL evaluation for twitter\_dev\_test.ner.pred(without features)**



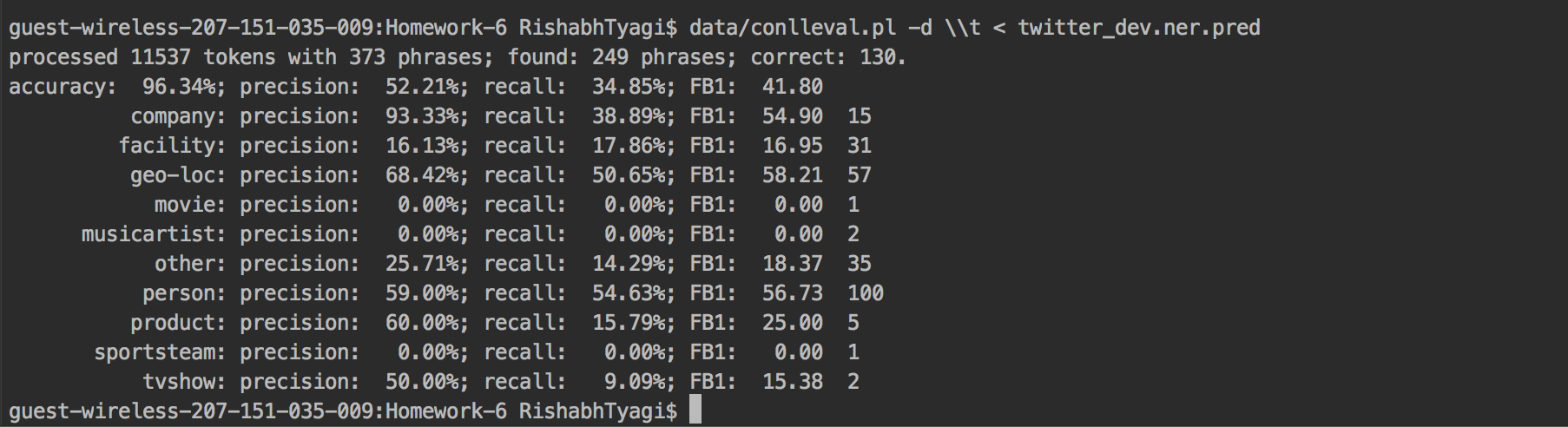
Accuracy of Logistic Regression: 91.02%; FB1: 13.51

Accuracy of Conditional Random Field: 91.31%; FB1: 23.84

**Graph for CRF vs LogReg for Twitter\_dev\_test.ner (without features):**

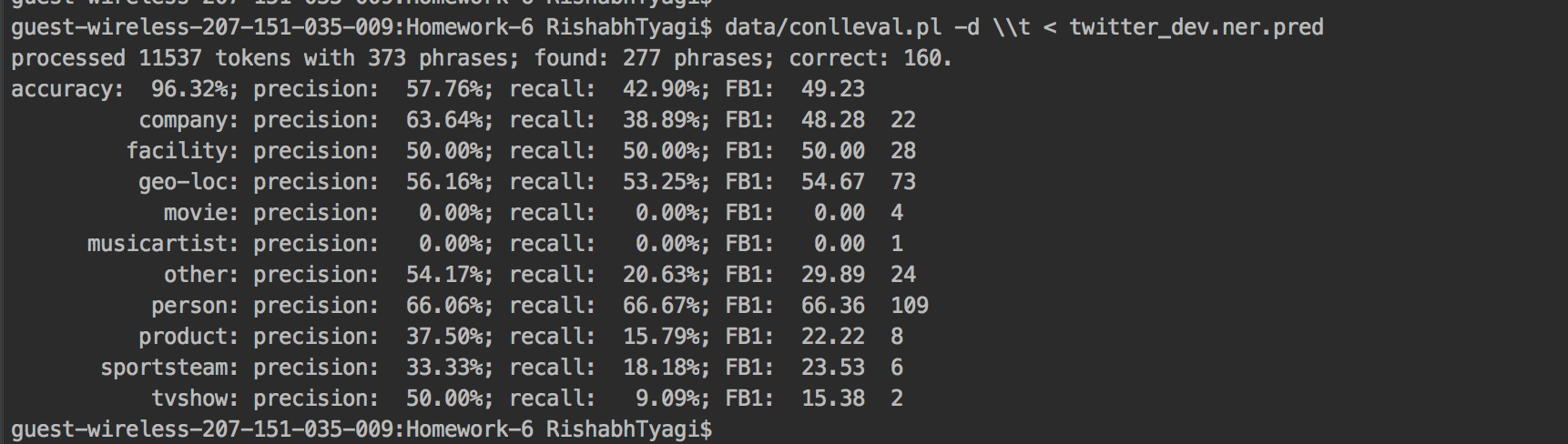


1. Comparing the difference in the output of the evaluation between logistic regression and conditional random field is shown below using baseline along with appended features.

**CONLL evaluation for twitter\_dev.ner.pred(with feature) : Using Logistic Regression Model**

**CRF: Conditional Random Field:**

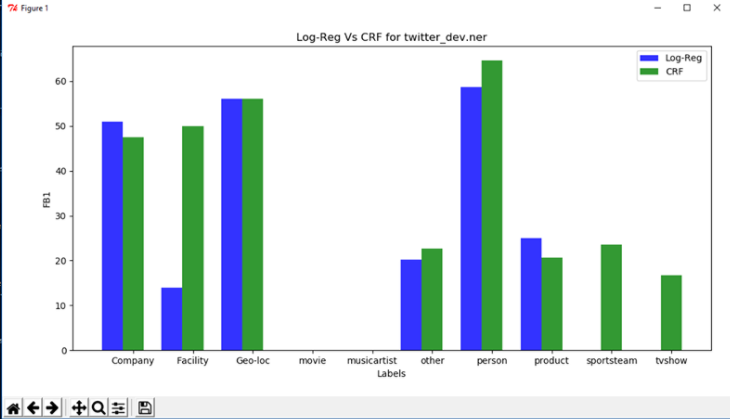
**CONLL evaluation for twitter\_dev.ner.pred(with features)**



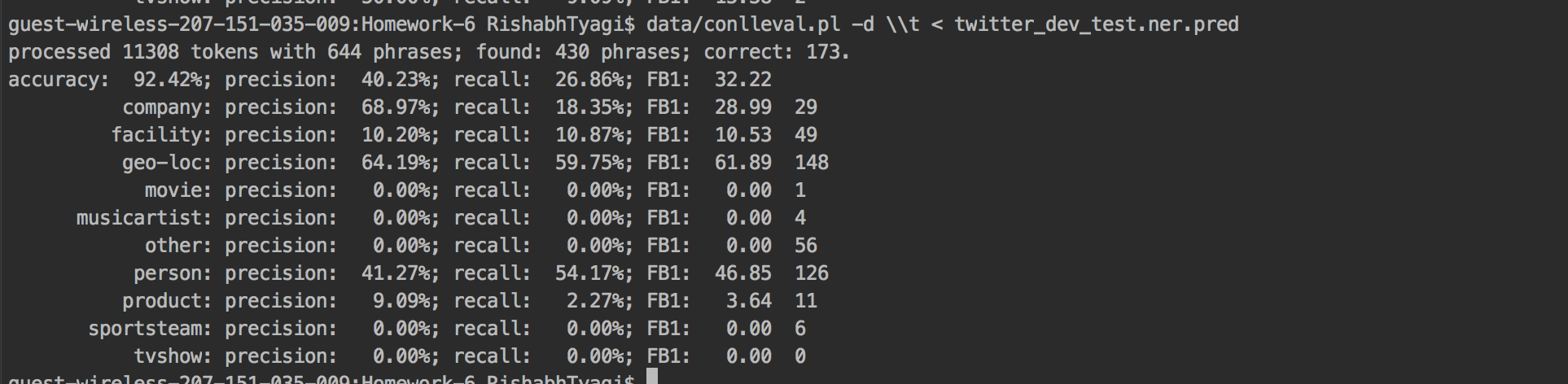
Accuracy of Logistic Regression: 96.34%; FB1: 41.80

Accuracy of Conditional Random Field: 96.32%; FB1: 49.23

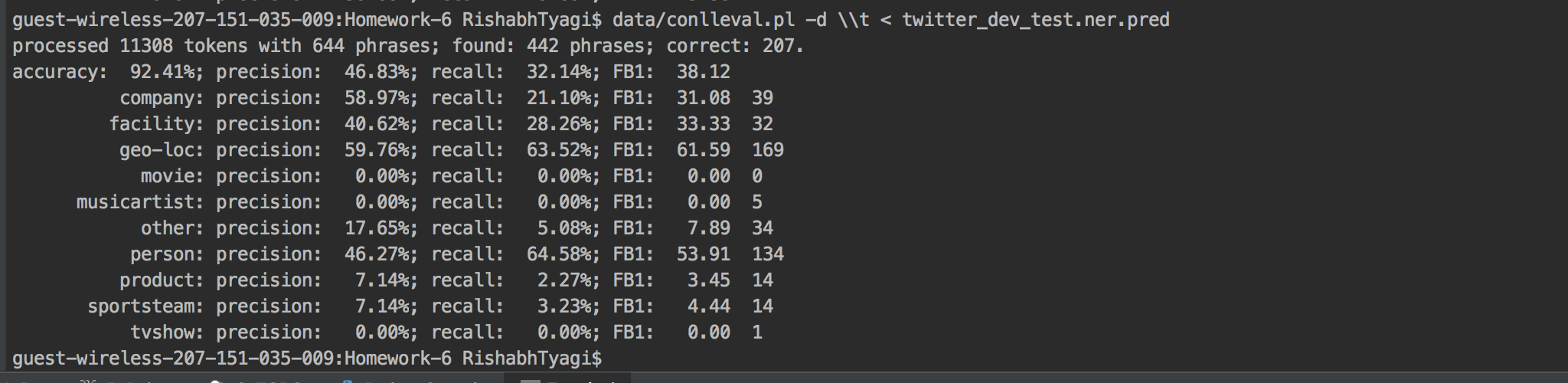
**Graph for CRF vs LogReg for twitter\_dev.ner (with features):**



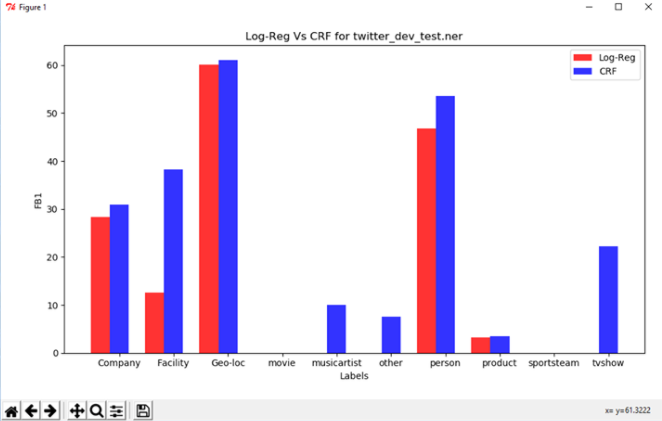
**CONLL evaluation for twitter\_dev\_test.ner.pred(with feature) : Logistic Regression Model**



**CONLL evaluation for twitter\_dev\_test.ner.pred(with features)**



**Graph for CRF vs LogReg for twitter\_dev\_test.ner (with features):**



The accuracy of both the model is almost same but the FB1 score is significantly larger in crf model which proves that crf is a better model as the number of false positives and false negative are significantly less in crf model. This is the reason why crf works better than logistic

As per me, crf model works better in sequence tagging problem, mainly because this model will account for the lables of the previous observed inputs as well the features of the current input using Viterbi algortihm. On the other hand, the logistic regression, will only take features of each token into account at a time, making it more susceptible to missing out on context of the word.