B Satya Kishore
2012CS50284
COL 772: Assignment 2 Writeup

Used CRF++
Used 2 level prediction
Used 17 features each spanning a context window of 7 words [-3 to 3]
Some features are
- isLink as the model is getting confused link with telephone number
- has9Digits
- numType which says if a number is big or small
- hasNum says if the word has number or not
- hasQuantifier says if it is 1st or 2nd or and so on
- isInAttributeList says if a part of word is in attribute gazetteers formed using the labeled data
- isInNameList says if a part of word is in attribute gazetteers formed using the data posted on piazza
- isPrice says if the word represents price
- isAreaUnit says if the word represents area
- isPer says if the word has an instance of per
- hasLocation says if a part of word is in location gazetteers formed using the data posted on piazza
- StartsWithAt says if the word starts with @
- startsWithHash says if the word starts with #
- isContact says if the word is either contact or call
- isIn says if the word is in or not
- isAt says if the word is at or not

Help Taken:
1. Kyasa Rohit - 2012MT50600
2. Shikhar for Names from Piazza