Medical Named Entity Recognition on Twitter Data

Group:56

Chinmay Bapna	Juhi Tandon	Kalpish Singhal
201302182	201225032	201505513

Index

- Introduction
- Approach
- Evaluation and Results
- Conclusion
- References

OUTLINE

- **Introduction**
- Approach
- Evaluation and Results
- Conclusion
- References

Introduction

- Medical Entity Recognition is a crucial step towards efficient medical texts analysis. *The task of a Medical Name Entity Recognizer is two fold. -*

    - Identification of entity boundaries in the sentences
    - Entity categorization.

  Our objective is to extend medical entity recognition for tweets.

- **Medical Entities**

  a. Diseases

  b. Drugs

  c. Symptoms

OUTLINE

- Introduction
- **Approach**
- Evaluation and Results
- Conclusion
- References

Approach

Tools Used:

1. Mallet:
   - We use Mallet for natural language processing,text classification.
   - The advantages of using Mallet is that it train CRFs with arbitrary graphical structure.
2. Meta-Map:

   - MetaMap is a highly configurable program developed to map biomedical text to the UMLS Metathesaurus.

3. Tweet NLP:

   - We used it to address the problem of part-of-speech tagging for English data from the popular microblogging service Twitter.

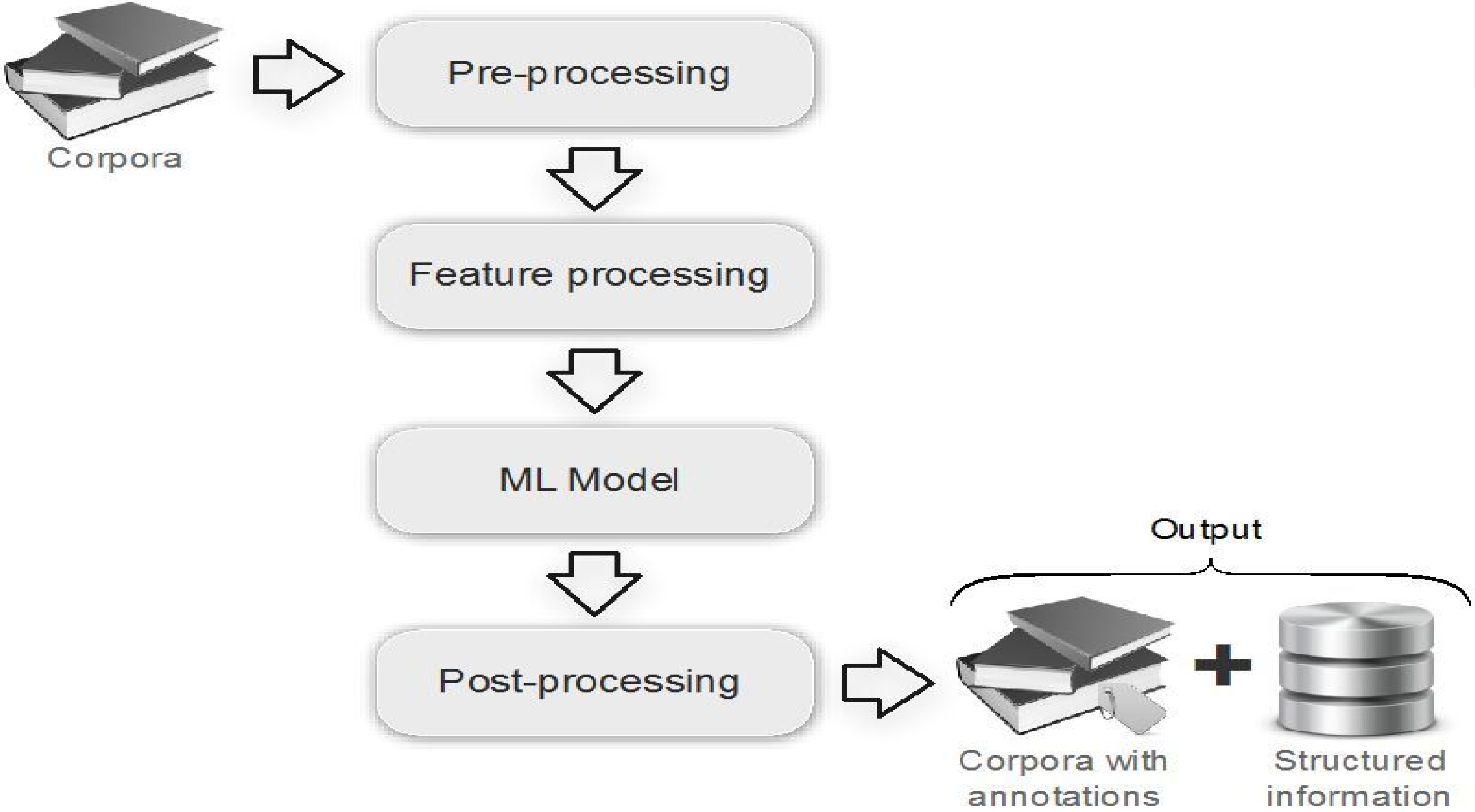
Workflow

1. Each tweet is stored in a separate file along with a corresponding annotation file with manually labelled entities
2. Next comes the word-context feature, along with the term, a 2-term context window, preceding and succeeding the concerned term were added as another feature
3. Metamap tagging: For each tweet, metamap tags are extracted
4. POS tagging: For each tweet, part of speech tags are extracted
5. Orthographic tagging: For each term, orthographic features are identified and labelled

Workflow

1. In Next step Brown-Clustering was used cluster combined data for all tweets from training and testing data.
2. Separate Feature Files were made for training and testing data, Mallet was trained on the feature file
3. From the Feature file created for Testing data, labels were extracted as a separate file, which would later be used for evaluation purposes.
4. The rest of the features from the Testing Feature file (all except labels) were fed to the tool Mallet as a testing file using the trained model constructed on training data..
5. The output tags obtained from the previous step -> were stored in a separate file, which would now be sent for evaluation.

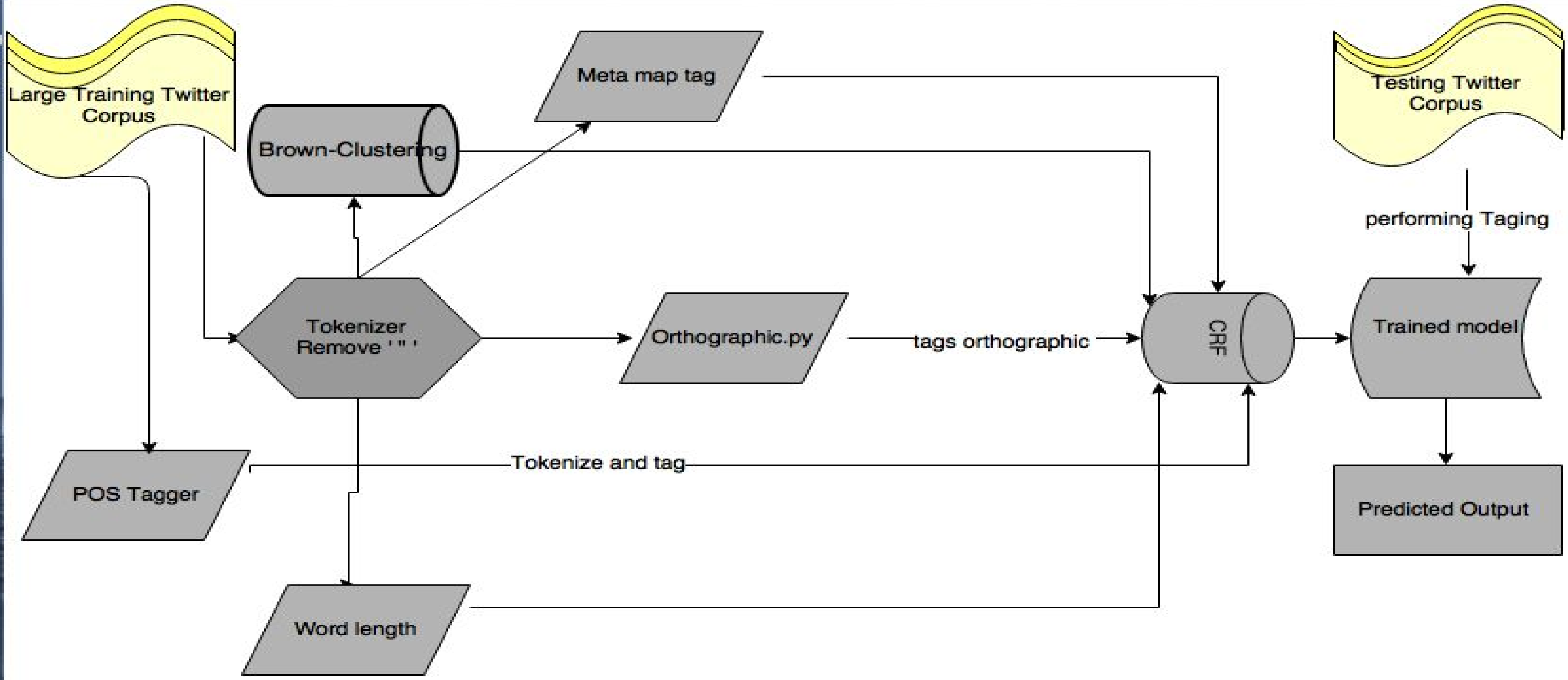
Overall Process



Features

Token	Meta-tag
PoS-tag	Average_word_length
Next-Noun next-Verb	Cluster-id
Ortho-tag	Term context
4_prefixes	4_suffix

System Architecture



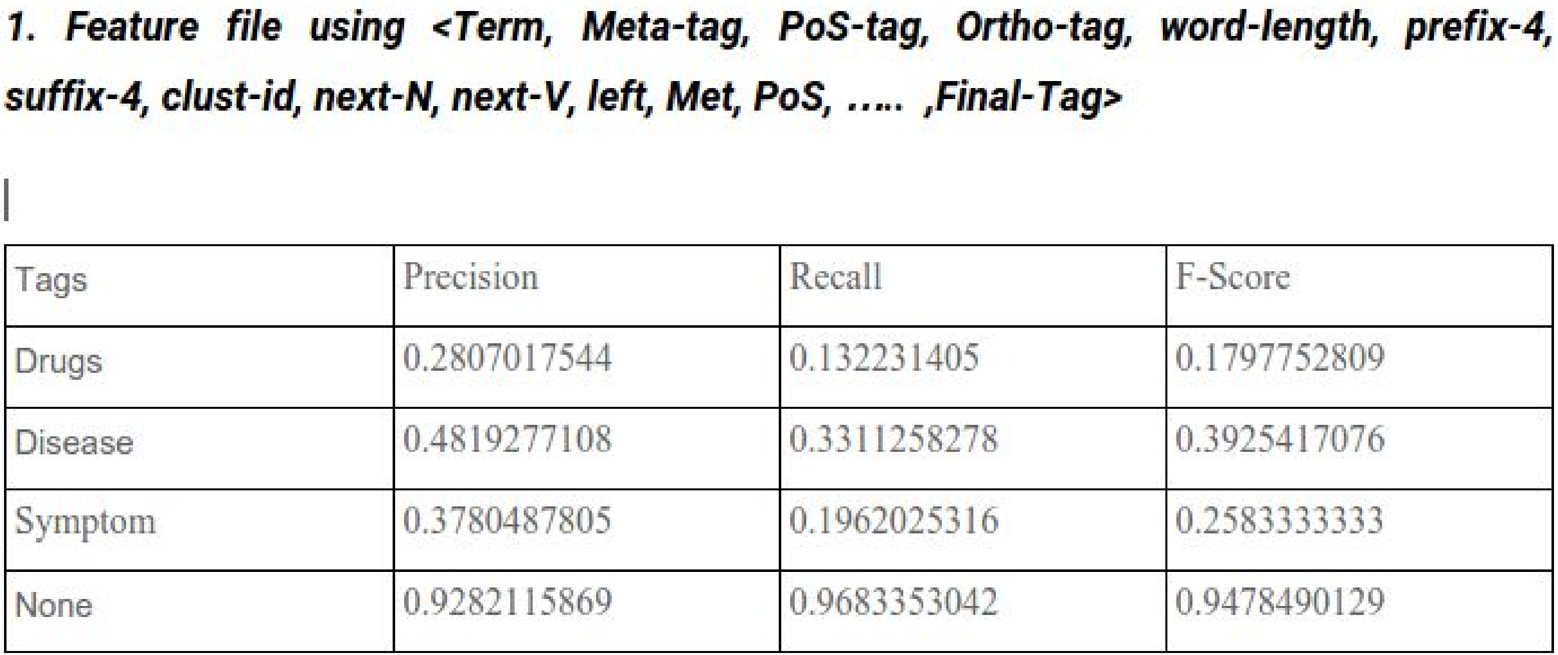
OUTLINE

- Introduction
- Approach
- **Evaluation and Results**
- Conclusion
- References

Evaluation and Results

- The two files obtained from the previous process, the system(mallet) generated tag file and the tag file generated from predefined annotations will be used for evaluating the efficiency in NER for considered set of features.
- 4 types of evaluation metrics
    1. Accuracy
    2. Precision
    3. Recall
    4. F-score

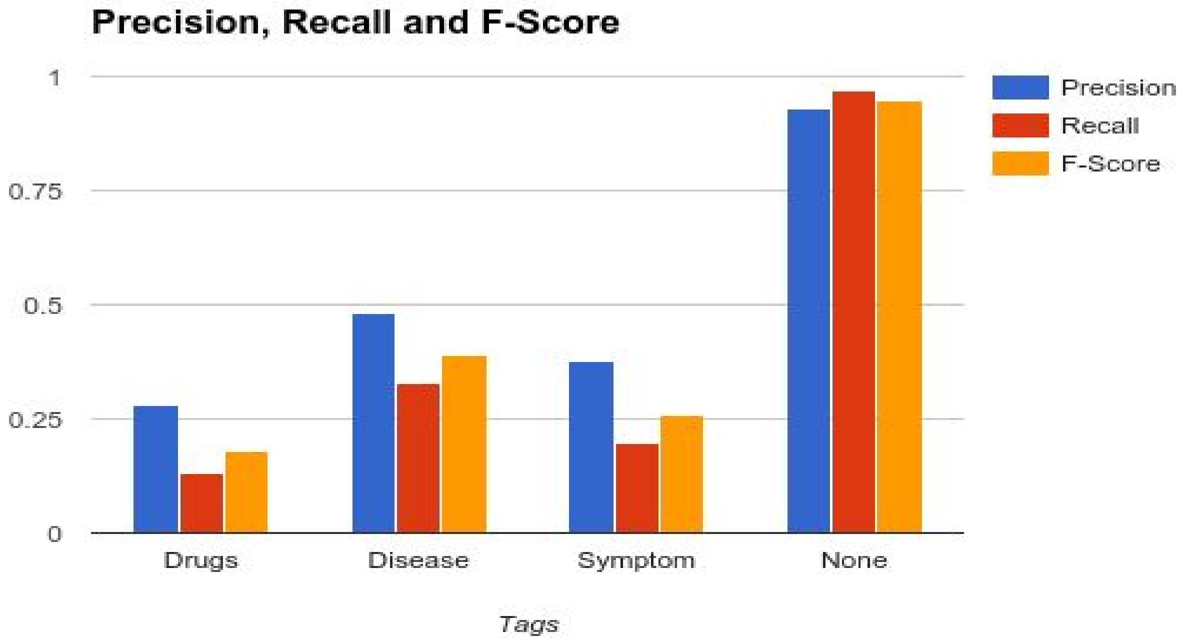
Results..



1. Feature file using <Term, Meta-tag, PoS-tag, Ortho-tag, word-length, prefix-4, suffix-4, clust-id, next-N, next-V, left, Met, PoS, ….. ,Final-Tag>

Tags	Precision	Recall	F-Score
Drugs	0.2807017544	0.132231405	0.1797752809
Disease	0.4819277108	0.3311258278	0.3925417076
Symptom	0.3780487805	0.1962025316	0.2583333333
None	0.9282115869	0.9683353042	0.9478490129

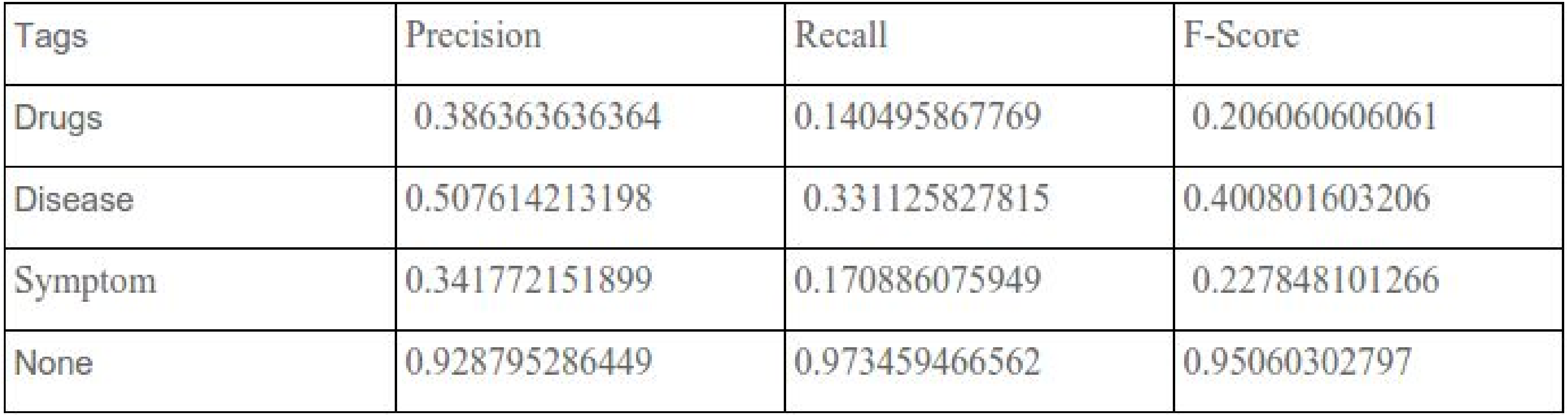
Results..



Accuracy =7617/8494=0.896750647516

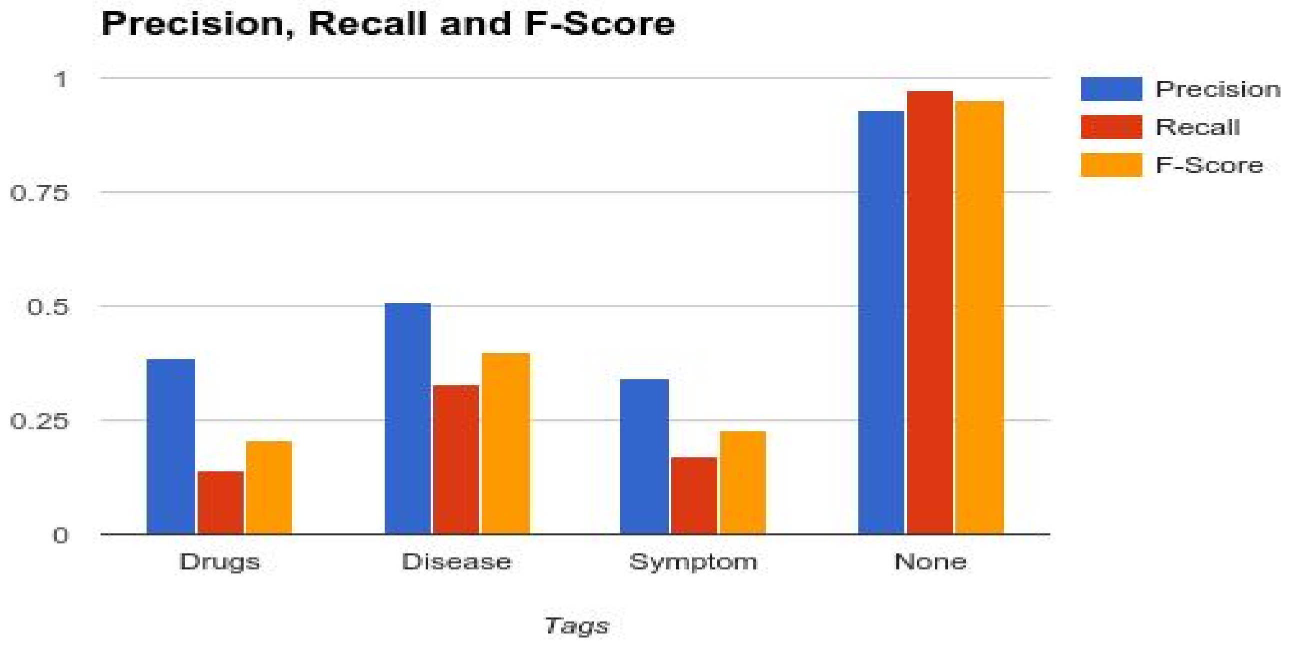
Results conti..

2. ***Term, PoS-tag, Ortho-tag, word-length, prefix-4, suffix-4, clust-id, next-N, next-V, left, PoS, ... ,Final-Tag***



Tags	Precision	Recall	F-Score
Drugs	0.386363636364	0.140495867769	0.206060606061
Disease	0.507614213198	0.331125827815	0.400801603206
Symptom	0.341772151899	0.170886075949	0.227848101266
None	0.928795286449	0.973459466562	0.95060302797

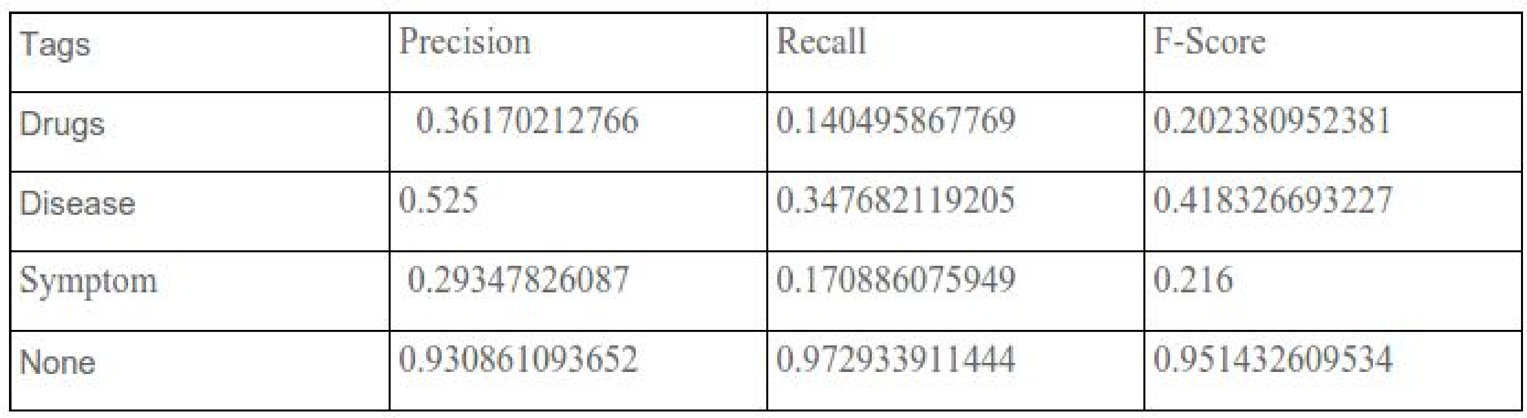
Results conti..



Accuracy =7653/8494=0.900988933365

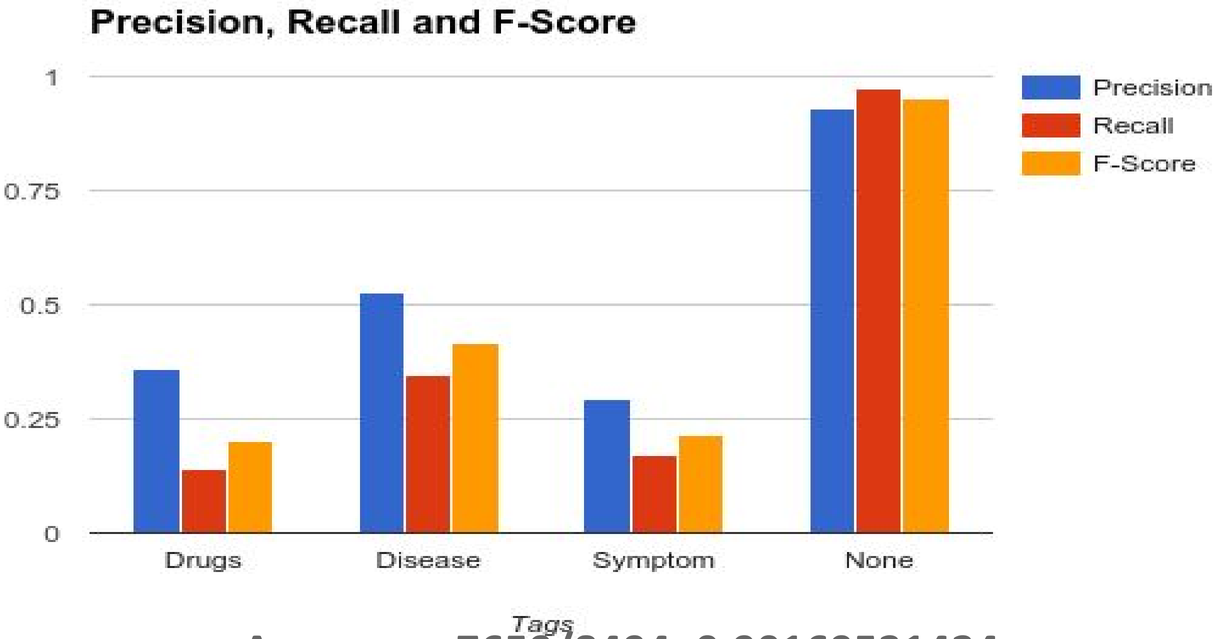
Results conti..

3.***Term, Ortho, Term-Context, Clust-id***



Tags	Precision	Recall	F-Score
Drugs	0.36170212766	0.140495867769	0.202380952381
Disease	0.525	0.347682119205	0.418326693227
Symptom	0.29347826087	0.170886075949	0.216
None	0.930861093652	0.972933911444	0.951432609534

Results conti..



Accuracy =7659/8494=0.90169531434

OUTLINE

- Introduction
- Related Work
- Approach
- Evaluation and Results
- **Conclusion**
- References

Conclusion

- The dataset provided to us consists of 85% of ‘None’ entity tags.
- Thus with any random designation of tags too we will have 80% or more accuracy
- This conveys that accuracy is not the correct metric for correct evaluation of feature models.
- Since this metric isn’t good-enough to tell us significance of one feature model over another,
- we tried following metrics:

    1. Precision,
    2. Recall
    3. F-Score

- In this project we experimented with different feature sets and analyzed their significance based on intuition and prior knowledge .
- Thereafter we evaluated their efficiency in Medical-NER.

OUTLINE

- Introduction
- Approach
- Evaluation and Results
- Conclusion
- **References**

References

- Medical Entity Recognition: A comparison of semantic and statistical methods http://www.aclweb.org/anthology/W11-0207
- Enhancing clinical concept extraction with distributional semantics http://ac.els-cdn.com/S1532046411001730/1-s2.0-S1532046411001730-main.pdf?_tid=0ee50780-0281-11e6-944d-00000aacb360&acdnat=1460666580_8a4179d1464230d82e1e8907cb09ad9b
- Using Twitter Data and Sentiment Analysis to Study Diseases Dynamics http://link.springer.com/chapter/10.1007%2F978-3-319-22741-2_2#page-1

Links

- Git web page :-http://kalpishs.github.io/IRE--Medical_NER_Twitter/
- Git Repo :- https://github.com/kalpishs/IRE--Medical_NER_Twitter
- Youtube:- https://youtu.be/dFKIy7CgMrg
- Dropbox link:- https://goo.gl/3Plc8s

THANK YOU!!