Negation Detection using NooJ

Gaurish Thakkar, Nives Mikelic Preradovic, Jeremy Barnes

Faculty of Humanities and Social Sciences in Zagreb
Croatia

Language Technology Group, University of Oslo
Norway

The project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 812997.

Outline

TASK OF NEGATION DETECTION

HOW DID WE SOLVE IT

WHAT WERE THE RESULTS

CONCLUSION & FUTURE WORK

Negation Detection

- Negation can be implicit, as in “with this act, it will be his first and last movie”—it carries a negative sentiment, but no negative words are used.
- Negation can be explicit, as in “this is not good.”
- Negation can be morphological where it is either denoted by a prefix (“dis-”,“non-”) or a suffix (“-less”).
  - "It seems a singularly useless thing to steal , '' said Sherlock Holmes .
  - I was not sure whether I had left it here or in the Shipping Office.

Problem Statement

- Given a sentence $S = \{ t_0 \dots t_n \}$ where t_i is token, the aim is to identify negation-cue $N_C = \{ t_i \dots t_k \}$ and negation-scope $N_S = \{ t_j \dots t_l \}$.

- Come , come , we are not so far wrong , after all , '' said Holmes .
  - (``,'O'),('Come','O'),(',','O'),('come','O'),(',','O'),('we','B_scope'),('are','I_scope'),('not','B_cue'),('so','B_scope'),('far','I_scope'),('wrong','I_scope'),(',','O'),('after','O'),('all','O'),(',','O'),("","O'),('said','O'),('Holmes','O'),('.','O)

Related Work

- Koza W. et al. (2018)
  - Negated findings in radiological reports. (medical terminology dictionary+ grammar rules).
- Tanushi, Hideyuki, et al. (2013)
  - Compared 3 different approaches for negation detection namely NegEx, PyConTextNLP and SynNeg.
- Chapman et al. (2001)
  - Negex regular expression-based algorithm.

Dataset

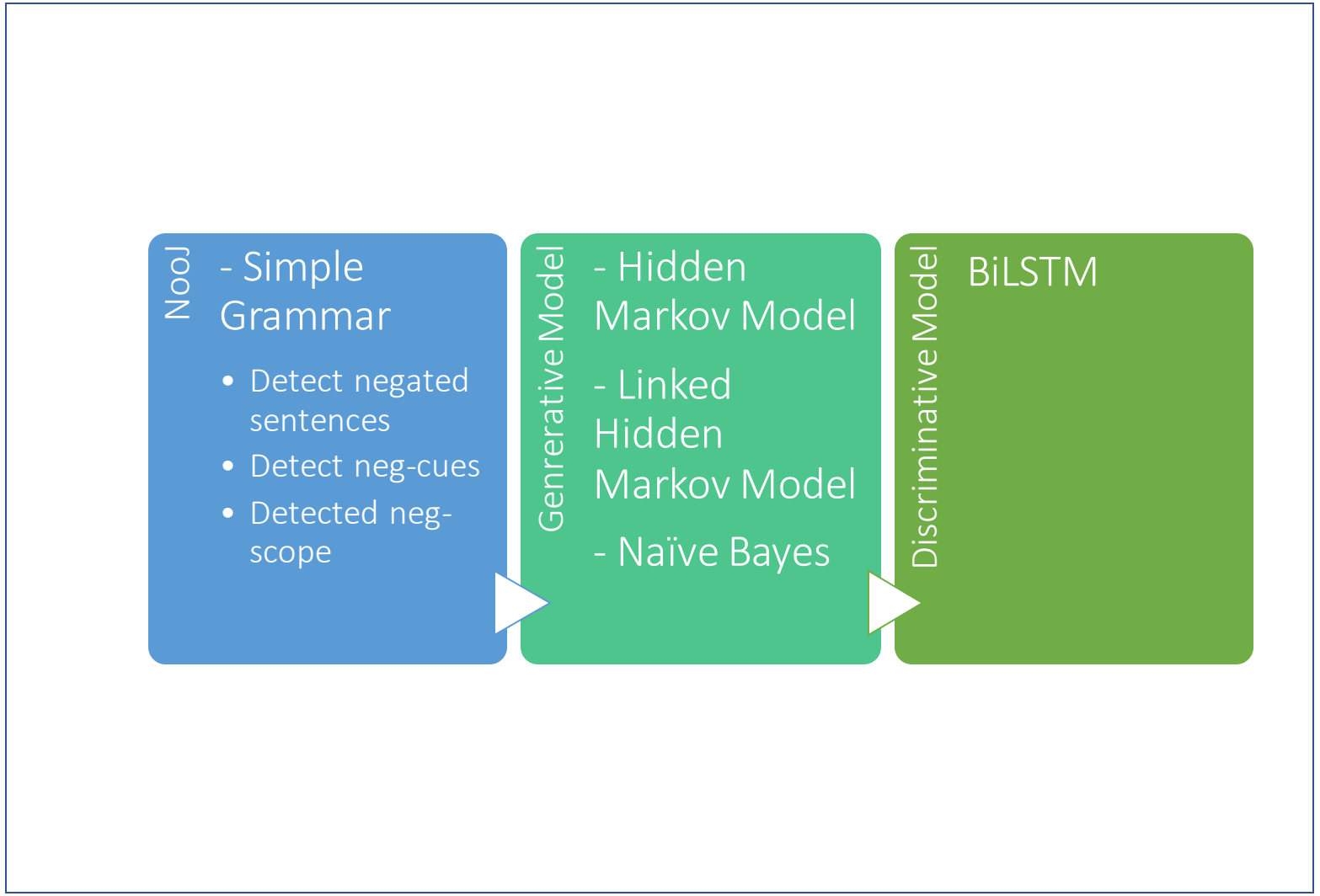
- The Conan Doyle neg-corpus

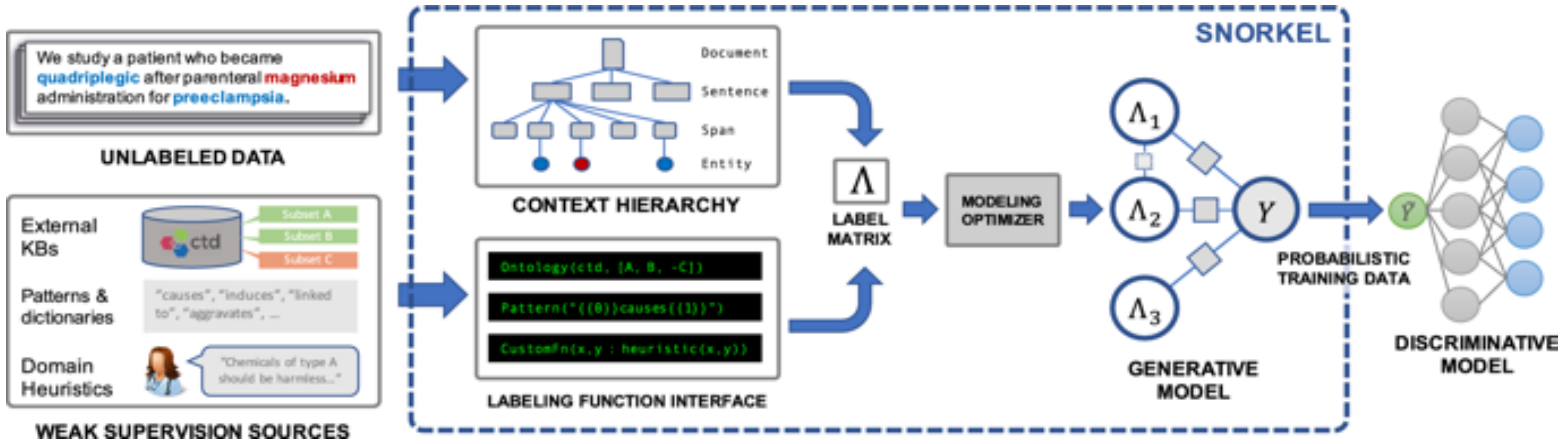
	Train	Dev	Test
Total	3644	787	1089
Negated	842	144	235
Number of sentences			

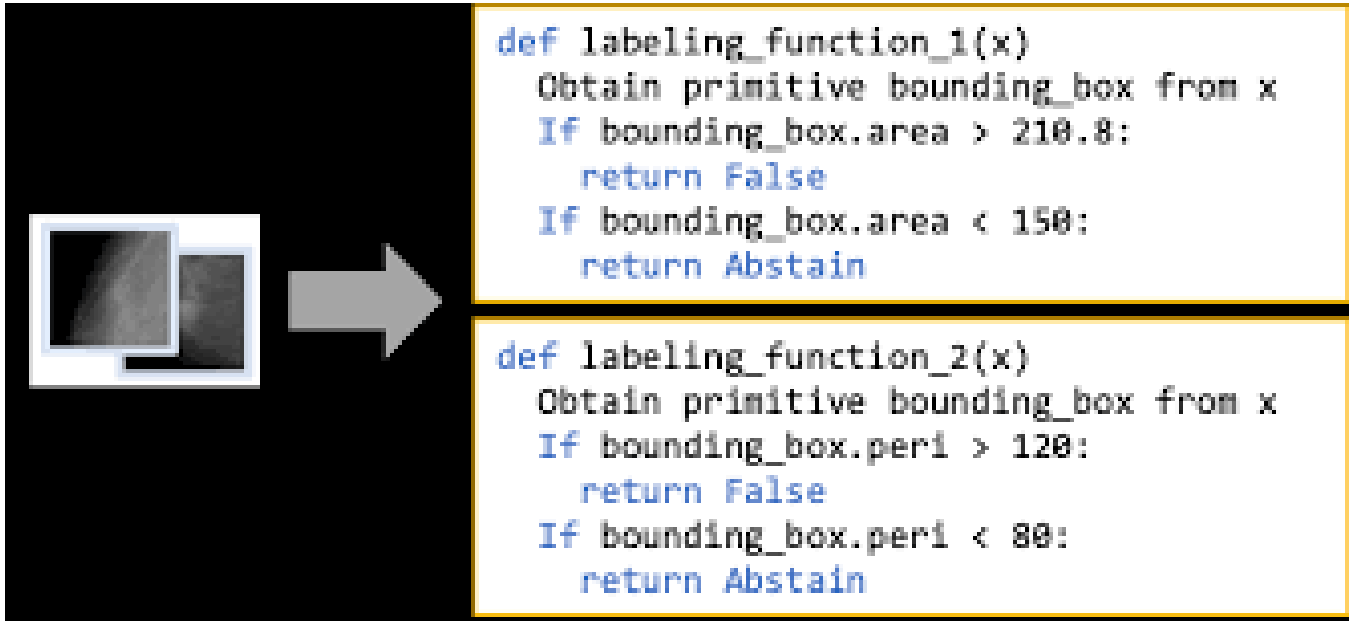
Experimental Setup

- Low-resourced setting
  - We do not use the annotated train set for supervision .
- Use dev set for tuning and test for computing final score.
- Unannotated version of train set is tagged via NooJ grammar is used for training subsequent systems.

Overall Diagram





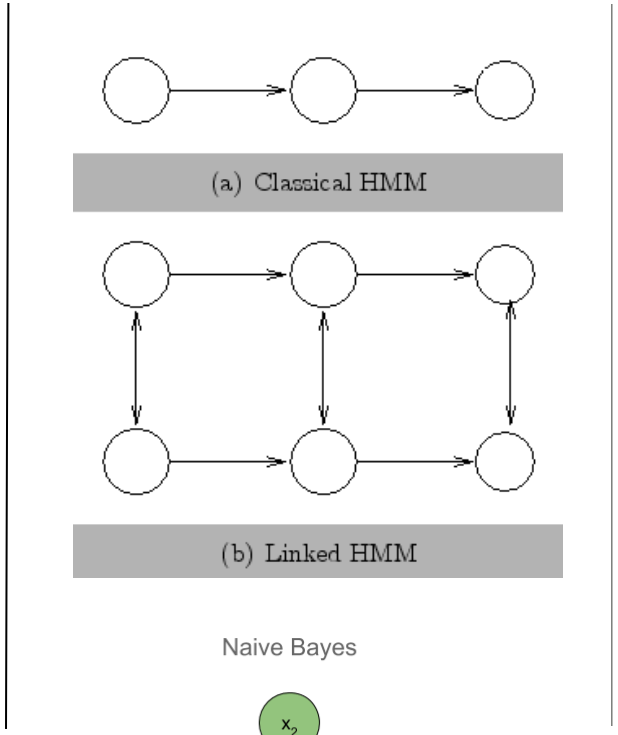


```
def labeling_function_1(x)
  Obtain primitive bounding_box from x
  If bounding_box.area > 210.8:
    return False
  If bounding_box.area < 150:
    return Abstain
```

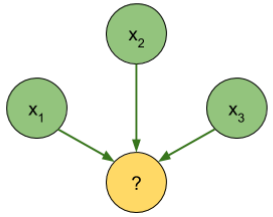
```
def labeling_function_2(x)
  Obtain primitive bounding_box from x
  If bounding_box.peri > 120:
    return False
  If bounding_box.peri < 80:
    return Abstain
```

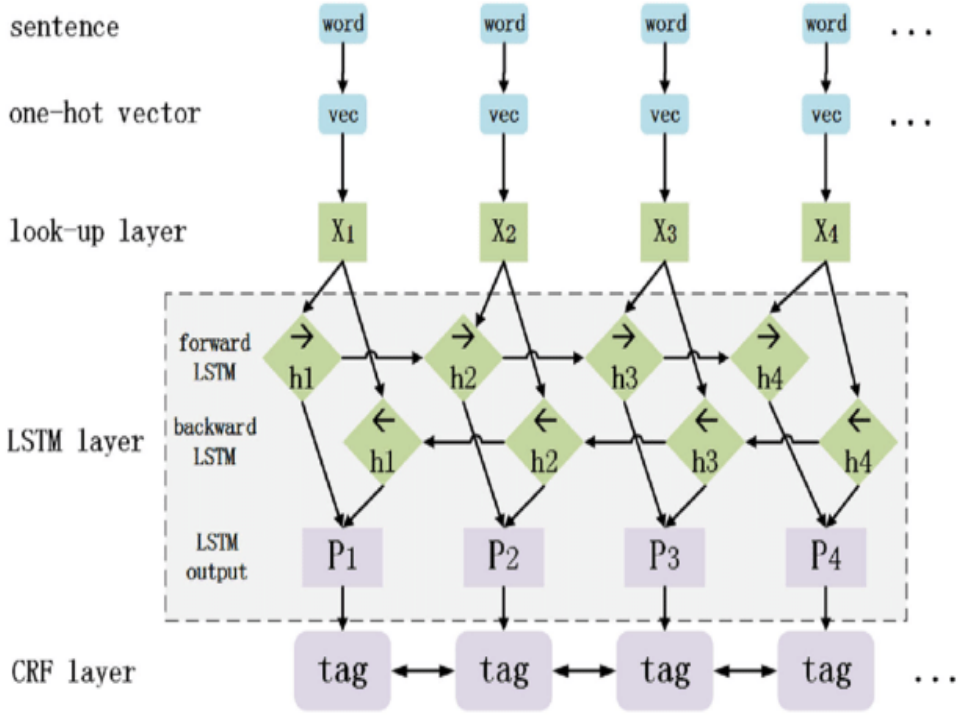
Data Programming

Models

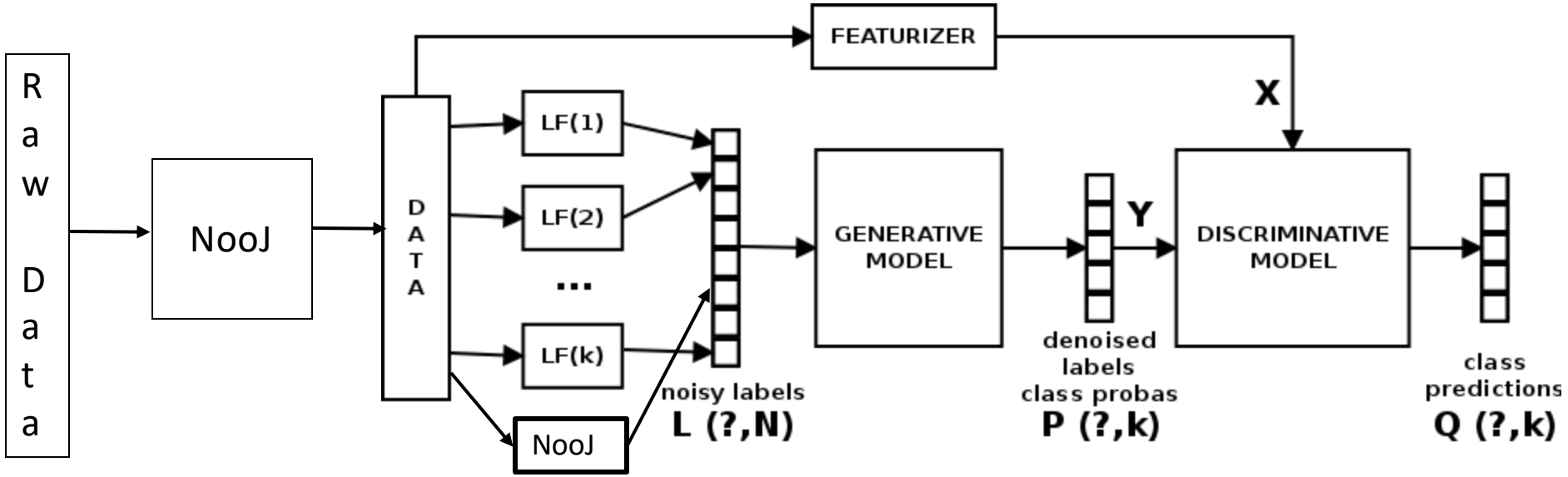


Naive Bayes

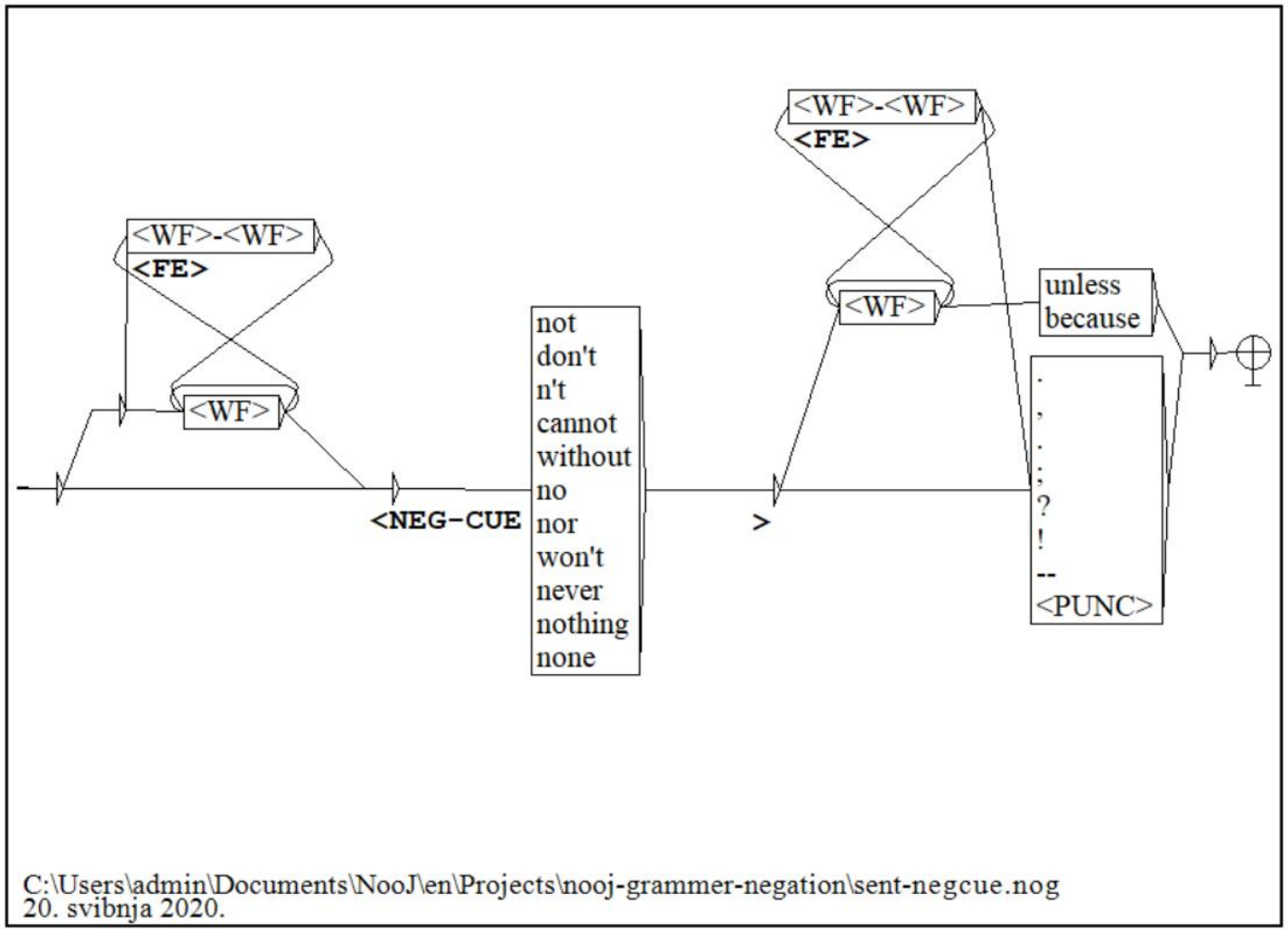




BiLSTM


Raw Data
NooJ
DATA
LF(1)
LF(2)
...
LF(k)
NooJ
noisy labels
L (?,N)
FEATURIZER
X
GENERATIVE MODEL
Y
denoised labels
class probas
P (?,k)
DISCRIMINATIVE MODEL
class predictions
Q (?,k)

Step 1: Nooj Grammar



Step 2: Labelling Functions

	TP	FP	FN	Token Acc.	Token Votes
BigramNegationCueNT	21	17	406	1.0000	38
CommonFalsePositives	0	0	427	1.0000	10
CommonTruePositivesImplicitsCue	140	5	287	0.9793	145
CommonTruePositivesImplicitsScope	1	10	426	1.0000	11
CueConstituencyKeywords	254	158	173	0.7948	1272
Non-EntityPunctuation	0	0	427	0.9511	348
NoojLabels	276	88	151	0.8033	2278
TrigramNegationCue	1	0	426	1.0000	3

Results

1. How many lines were detected using the simple Nooj Grammar?

- EXPT 1.1

	Detected	Total	Acc
Train	700	842	0.83
Dev	113	144	0.78
Test	202	235	0.85

2. What is the Precision, Recall and F1-score on applying the grammar for detecting the negation cue and scope?

	Precision	Recall	F1	Acc
Train	0.76	0.74	0.75	0.81
Dev	0.76	0.73	0.74	0.77
Test	0.75	0.73	0.74	0.82

**All score are Micro Avg

Results

Method	Precision	Recall	F-1
NooJ	-	-	-
NooJ+LinkedHMM+BiLSTM	0.75	0.68	0.71
NooJ+LF+LinkedHMM+BiLSTM	**0.76**	**0.73**	**0.74**
NooJ+LF+HMM+BiLSTM	0.75	0.73	0.74
NooJ+LF+NB+BiLSTM	0.53	0.74	0.62
LF+LinkedHMM+BiLSTM	0.**55**	0.71	0.62
Step 2: Generative Model			

Method	Precision	Recall	F-1
NooJ	0.75	0.73	0.74
NooJ+LinkedHMM+BiLSTM	0.75	0.73	0.74
NooJ+LF+LinkedHMM+BiLSTM	**0.76**	**0.75**	**0.76**
NooJ+LF+HMM+BiLSTM	0.75	0.73	0.74
NooJ+LF+NB+BiLSTM	0.67	0.74	0.70
LF+LinkedHMM+BiLSTM	0.50	0.74	0.60
Step 3: Discriminative Model			

Error Analysis

- A lot of false positives are introduced e.g. but, although
- Current setup fails to capture complex sentences
  - # Mr. Sherlock Holmes , who was usually very late in the mornings , save upon those not infrequent occasions when he was up all night , was seated at the breakfast table .
  - # `` Now , tell me , Dr. Mortimer -- and this is important -- the marks which you saw were on the path and not on the grass ? ''

Conclusion & Future Work

- Simple grammar graphs do perform well.
- NooJ can be leveraged as a labelling function for training NLP systems.
- Improve grammar
  - Handle false positives
  - to handle long-range dependencies.
- Repeat the experiment with other languages with low-resources.
- Utilize the setup for other NLP tasks.

References

- Ratner, Alexander J., et al. "Data programming: Creating large training sets, quickly." *Advances in neural information processing systems*. 2016.
- Ratner, Alexander, et al. "Snorkel: Rapid training data creation with weak supervision." *The VLDB Journal* 29.2 (2020): 709-730
- Safranchik, Esteban, et al. "Weakly Supervised Sequence Tagging from Noisy Rules." *Thirty-Fourth AAAI Conference on Artificial Intelligence*. 2020.
- Koza W. et al. (2018) Automatic Detection of Negated Findings with NooJ: First Results. In: Silberztein M., Atigui F., Kornyshova E., Métais E., Meziane F. (eds) Natural Language Processing and Information Systems. NLDB 2018. Lecture Notes in Computer Science, vol 10859. Springer, Cham
- Tanushi, Hideyuki, et al. "Negation scope delimitation in clinical text using three approaches: NegEx, PyConTextNLP and SynNeg." 19th Nordic Conference of Computational Linguistics (NODALIDA 2013), May 22-24, 2013, Oslo, Norway. Linköping University Electronic Press, 2013.
- Chapman, W. W., Bridewell, W., Hanbury, P., Cooper, G. F., and Buchanan, B. G. (2001b). A simple algorithm for identifying negated findings and diseases in discharge summaries.JBiomed Inform, pages 34–301.

Questions ?