Improving Distantly Supervised Relation Extraction using Word and Entity Based Attention

Sharmistha Jat*, Siddhesh Khandelwal* and Partha Talukdar

{sharmisthaj,ppt}@iisc.ac.in,skhandel@cs.ubc.ca

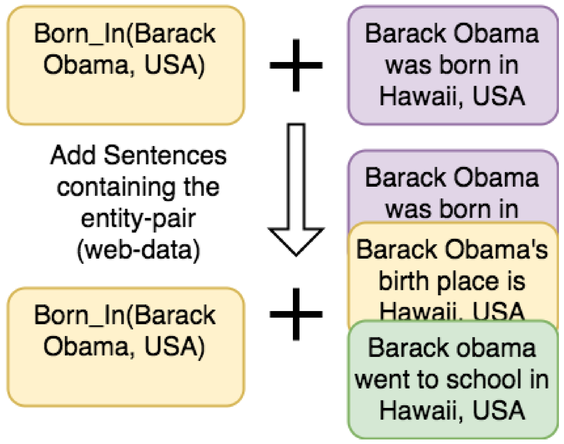
Contributions

Our main contributions are:

1. We introduce the Google-IISc Distant Supervision (GIDS) dataset, a new dataset for distantly-supervised relation extraction
2. We propose two novel word attention based models for distant supervision, viz., BGWA, a BiGRU-based word attention model, and EA, an entity-centric attention model
3. We show efficacy of combining new and existing relation extraction models using a weighted ENSEMBLE model. Weights of each model learnt using linear regression

GIDS: New Dataset for Distant Supervision

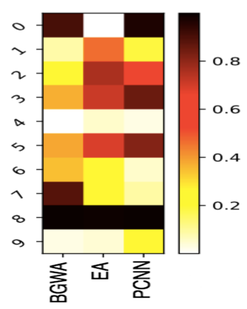
In the new Google-IISc Distant Supervision(GIDS) dataset, each sentence-set for an entity-pair is guaranteed to contain at least one sentence which expresses the relation assigned to that set. Each relation is human verified helping in credible test dataset for evaluation.



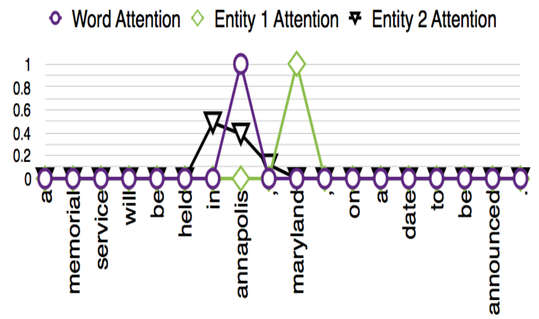
Relations	#sentence	#entitypair
perGraduatedInstitute	4456	2624
perHasDegree	2969	1434
perPlaceOfBirth	3356	2159
perPlaceOfDeath	3469	1948
NA	4574	2667

Dataset	#relation	#sentences	#entity-pair
Train	5	11297	6498
Dev	5	1864	1082
Test	5	5663	3247

Why Attention and Ensemble?



Confidence scores (indicated by color intensity, darker is better) of models on true labels of 10 randomly sampled instance sets from GIDS.



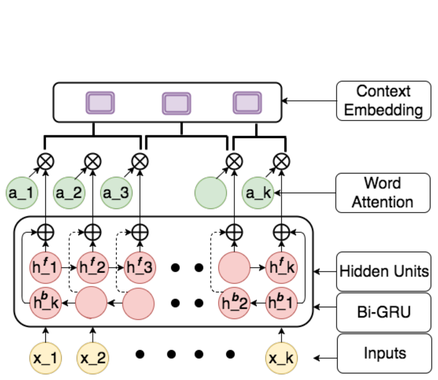
BGWA (Word attention) and EA (Entity Attention) values for an example sentence between entity pair *maryland* and *annapolis* and relation *location_in*. X-axis shows the sentence words and y-axis shows the attention scores.

WHY new dataset? *Noise* in Distantly supervised dataset

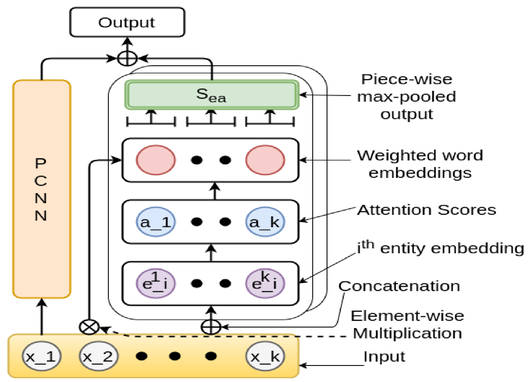
Examples of Noise in dataset. Sample 1,2 are incorrectly labelled with NA relation in the test set due to missing facts in Knowledge Base (KB). While, Sample 3's single sentence in the instance set does not support the KB relation.

Entity 1	Entity 2	Test Set Label	Classified Relation
1. Marlborough	New Hampshire	NA	/location/location/contains
2. Katie Couric	CBS	NA	/business/person/company
Entity 1	**Entity 2**	**Test Set Label**	**Instance Set**
3. Gary Sheffield	Florida	/people/person/place_lived	others who have already indicated they will wearno. 42 include ken griffey jr. of cincinnati, florida's dontrelle willis, carlos lee of houston, lee of the cubs and detroit's gary_sheffield .

Word Attention Models



BGWA (Word attention)



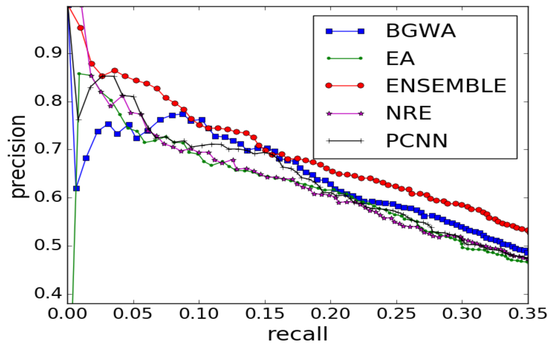
EA (Entity Attention)

*Equal Contribution Authors

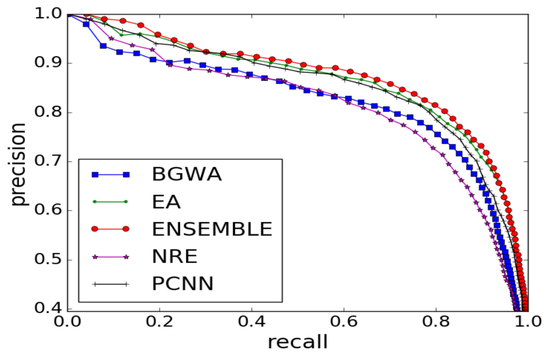
Results

Results Highlights:

1. **Word Attention** helps the model focus on relevant part of a long sentence to perform better on the relation extraction task.
2. **Ensemble method** for relation extraction works well due to model complimentarity
3. **New GIDS dataset available** for evaluating distantly supervised relation extraction methods without human verification of labels.



Precision-recall curves of our models against traditional state of the art methods for Riedel2010-b dataset with development set. We partition Reidel2010-a's train set into a new train (80%)and development set (20%). Development set if created to facilitate the learning of an ensemble model.



Precision-recall curves of our models on GIDS dataset

Acknowledgement

This work was supported by grants from Google, Accenture, and Government of India.