Named Entity Disambiguation with Graphs

https://www.youtube.com/watch?v=f_MCnUBWJCk

Sruthi Poddutur
{poddutur}@Stanford.edu

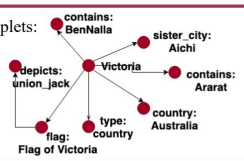
1. Problem and DataSet

- Named Entity Disambiguation removes ambiguity between mentions in the text vs referred entities in our Knowledgebase
- WikidataDisamb(Sloan et al 2018) dataset
  - derived from Wiki-disamb30
- Training size 250,000 and Test 10,000

2. SampleData

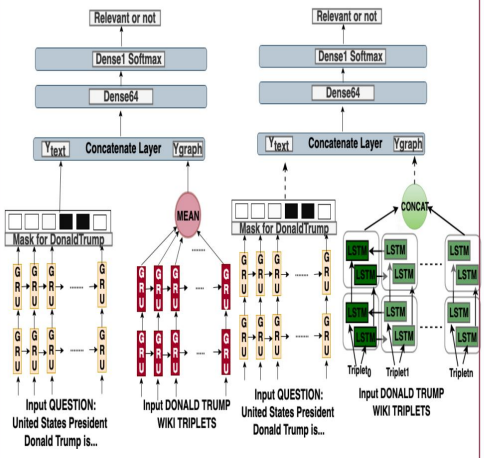
Entity	CorrectId	WrongId	Question
Victoria	Q36687	Q7926536	Fitzroy North is a suburb in Melbourne, Victoria, Australia, 4 km north-east from Melbourne's central business district Its Local Government Area are the Cities of Yarra and Moreland.

Q36687's Sample Triplets:



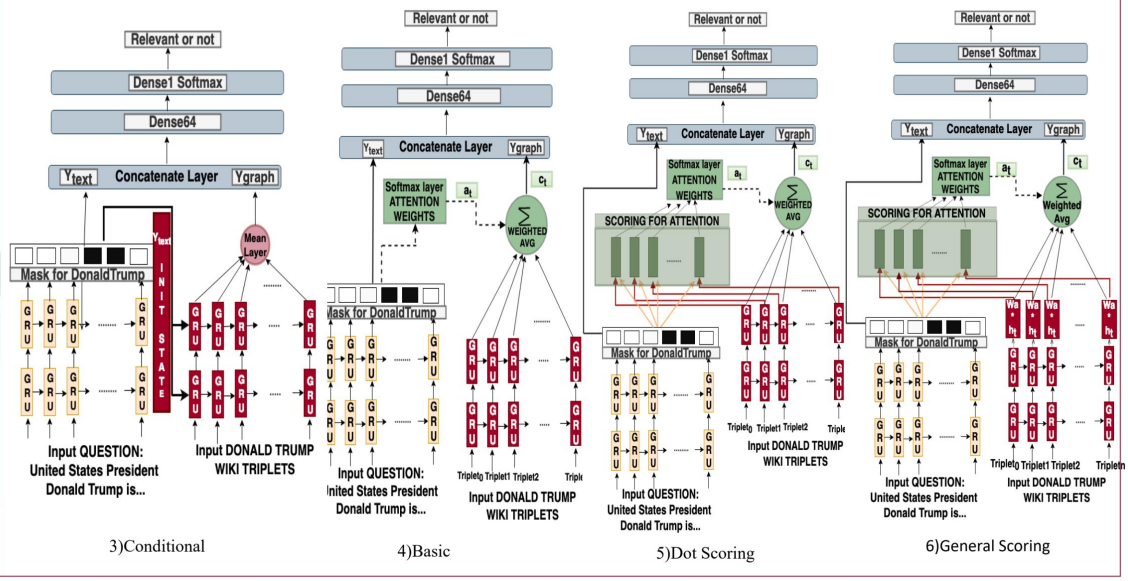
4. Models

RNN Models



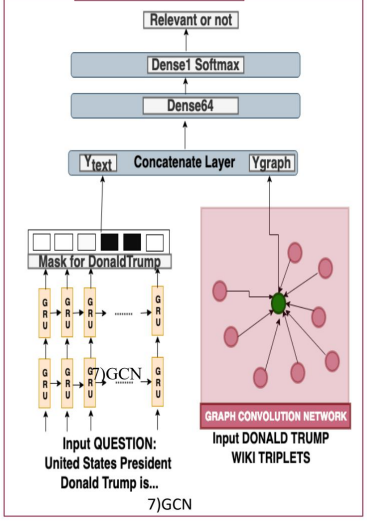
1)Mean triplet aggregation 2)Bi-LSTM triplet learning

RNN + Attention Models



3)Conditional 4)Basic 5)Dot Scoring 6)General Scoring

RNN + GCN Model



7)GCN

5. Results

Experiment Results on test dataset for various models

	positive records			negative records		
Description	**prec**	**rec**	**f1**	**prec**	**rec**	**f1**
GRU for both Y_{text} and Y_{text}	0.77	0.73	0.75	0.74	0.78	0.76
GRU for Y_{text} and Bi-LSTM for Y_{text}	0.80	0.76	0.78	0.77	0.81	0.79
RNN + Conditional Attention	0.84	0.82	0.83	0.83	0.84	0.83
RNN+Basic Attention	0.80	0.81	0.80	0.81	0.80	0.80
RNN+Attention on dot scoring function	0.88	0.88	0.88	0.88	0.88	0.88
RNN+Attention on general scoring function	0.90	0.89	0.89	0.89	0.90	0.90
GCN Based Model for Y_{text}	0.83	0.85	0.84	0.85	0.82	0.84

6. Conclusion

- Attention models performed better than basic RNN models matching what Cetoli et al 2018 observed
- Attention with General Scoring performed the best and expectedly performs better than basic attention
- GCN with one hop was comparable to other attention based models
  - improved results compared to Cetoli et al 2018 can be attributed to 1) addition of self or Identity to GCN computation and 2) Node degree normalization

Acknowledgements

Would like to acknowledge teaching staff and especially my guide Sagar Honnugar for guidance and discussions

7. Future Work

- GCN with multiple hops
- GCN with attention
- Various GCN message passing
- Use the entire 2 million data set
- Make an app given text gives Wiki ids of all entities
- Attention with triplet masking

8. References

- Cetoli 2018 arxiv.org/pdf/1810.09164.pdf
- Kipf 2017 arxiv.org/pdf/1609.02907.pdf

*Minh-thang 2015 arxiv.org/pdf/1508.04025.pdf