

Combining Contextual Words and Knowledge

Graph Embeddings

Software project, work update 4

IDMC, University of Lorraine

24/11/2020

Team members :

1. Rameez Qureshi
2. Vo Tuan Anh
3. Fatima HABIB
4. Asmaa Demny
5. Minh Houg Ngo

OUTLINES

1. Recap.
2. Current state : Auto-encoders Meta Embeddings
3. Future work

Recap

Dimensionality Reduction using PCA

Relation Prediction

Original Models (Dimension)	MRR	MAP@k=1
Contextual Embeddings (2048)	0.554	0.554
KG Embeddings (800)	0.817	0.663
Concatenation (2848)	0.738	0.533

Reduced Dimension Models (Dimension = 400)	MRR	MAP@k=1
Contextual Embeddings	0.656	0.433
KG Embeddings	0.750	0.557
Concatenation	0.708	0.481

Entity Typing

Original Models (Dimension)	MAP@k=10	Precision @ k (mean \pm sd)
Contextual Embeddings (1024)	0.631	0.449 \pm 0.271
KG Embeddings (400)	0.825	0.528 \pm 0.269
Concatenation (1424)	0.828	0.527 \pm 0.268

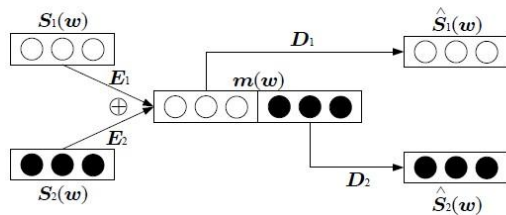
Reduced Dimension Models (Dimension = 200)	MAP@k=10	Precision @ k (mean \pm sd)
Contextual Embeddings	0.190	0.208 \pm 0.249
KG Embeddings	0.673	0.476 \pm 0.261
Concatenation	0.297	0.306 \pm 0.277

Relation Prediction with Concatenated Model

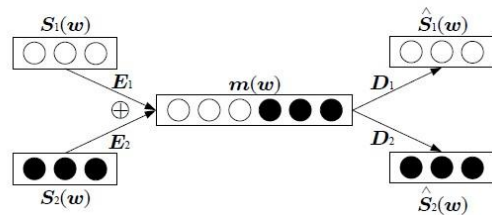
Dimension	MRR	MAP@k=1
400	0.708	0.481
800	0.716	0.493
1200	0.710	0.476
2848	0.738	0.533

Current state : Auto-Encoder Meta-Embedding

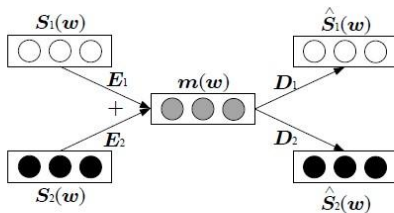
Decoupled, Concatenated and Averaged AEMEs



(a) DAEME



(b) CAEME



(c) AAEME

Training Meta Embeddings using AutoEncoders

- Input should be token-embedding matrix.
- Not possible with Elmo, as it requires context with the tokens.
- Required to find a unique common entity between graph embeddings and Elmo.
- Therefore, Knowledge Base IDs has been used instead of word token as unique ID.

Results :

- The trained model gives unexpected results: this may be caused by an error in the model structure or training process.
- Still working on figuring out the reasons behind these results .

Model	MAP@k	Mean Precision@k	Std
AAEME	0.101	0.200	0.282
CAEME	0.137	0.246	0.299
DAEME	0.125	0.198	0.212

Future work

To do :

- Before next meeting (in parallel) :
 - a. Re-produce the Auto-encoder model in “Learning Word Meta-Embeddings by Autoencoding” inorder to figure out the reason behind the low results.
 - b. Implementing the multihead attention mechanism.

References :

1. Bollegala, Danushka, and Cong Bao. "Learning word meta-embeddings by autoencoding." Proceedings of the 27th International Conference on Computational Linguistics. 2018.
2. James O'Neill and Danushka Bollegala: Semi-Supervised Multi-Task Word Embeddings arXiv, 2018. [arXiv]

Thank you! Questions?

