



Syntax Aware Word Embeddings Using Equivariant Graph Neural Networks

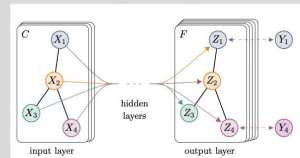
Pratham Soni

CS 224N: Natural Language Processing with Deep Learning, Winter 2022

Word embeddings are important

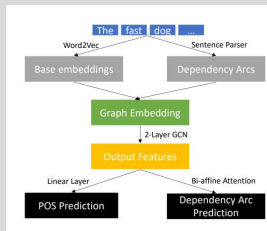
- Often static, do not respect sentence context (say Word2Vec)
- Contextual embeddings, like BERT, have been successful
- Syntactically Informed Word Representations (SIWRs) are learned using lexical information

Learning language From graphs

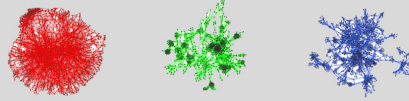


- GCNs pool node information from neighbors

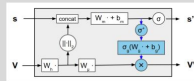
- Construct graph from base embeddings and syntactic parsing
- Pass through GCN to predict POS, dependency arcs at each node
- Serves as auto-encoder



Taking inspiration from proteins



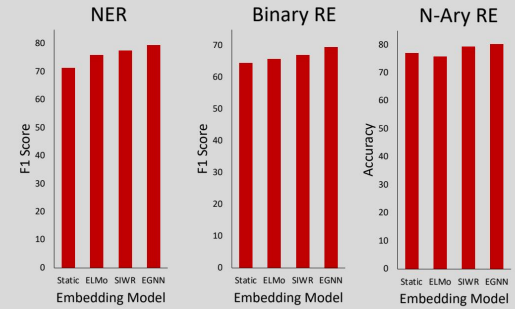
- Modelling proteins as graphs with vector features has been highly successful
- AI + protein studies have taken inspiration from NLP; can we flip the script?
- **Equivariant Graph Neural Network** replaces regular GCN from SIWR paper
- EGNN updates combine scalar and vector info



Embeddings in downstream tasks

- **Nested Entity Recognition:** Predict entity labels for words (ACE 2005, Wikipedia embeddings)
- **Binary Relationship Extraction:** Predict relationship between pairs of entities in a sentence (ACE 2005, Wikipedia embeddings)
- **N-Ary Relationship Extraction:** Predict relationship between n entities across multiple sentences (Drug Gene Mutation, PubMed embeddings)

Results



Discussion

- EGNN based embeddings improve performance in down-stream tasks
- Relatively light-weight model with only 1.5M additional params (BERT has >300 M)
- Potential to improve large LM embeddings (BERT/ELMo/GPT)

Acknowledgements & References

- My TA mentor, Angelica Sun, and the rest of the teaching team
- <https://arxiv.org/pdf/1609.02907v4.pdf> (GCN image)
- http://interactome.dfs.harvard.edu/S_cerevisiae/ (protein images)
- Thy Thy Tran, Makoto Miwa, and Sophia Ananiadou. Syntactically-informed word representations from graph neural network. Neurocomputing, 413:431–443, 2020.
- Bowen Jing, Stephan Eismann, Pratham N. Soni, and Ron O. Dror. Equivariant graph neural networks for 3d macromolecular structure. ArXiv, abs/2106.03843, 2021.