

Research Interests

Sebastian Ruder

May 6, 2016

My current work focuses mostly on developing new Deep Learning-based algorithms for sentiment analysis. Generally, I am interested in all aspects of Deep Learning for NLP. Concretely, I am most excited about the following three research directions:

1 Representation learning

As much of learning depends on finding the best mental images or representations for concepts, representation learning is a field that tremendously interests me. [Mikolov et al., 2013] and [Le and Mikolov, 2014] have both shaped my initial interest in this domain. In this context, I am currently working on leveraging bytes to pre-train cross-lingual word embeddings and on generating better aspect representations for aspect-based sentiment analysis. While I am interested in creating better (cross-lingual) word and document representations, for future work, I would like to find better ways to represent other forms of latent information, e.g. entities, discourse structures, speakers, world knowledge, etc. Inspiring works to this effect are [Gupta et al., 2015] who show that word embeddings partially encode entity attributes; [Li et al., 2016] who encode the persona and addressee of conversational agents; and [Iyyer et al., 2016] who represent relationships between book characters. For other latent factors such as emotions, psychological research has produced well-established representation taxonomies; finding ways to ground an expressive model in such existing mental models or leveraging these to inform an architecture despite the lack of quality data would be another interesting challenge.

2 Compositionality

Much of language is compositional and finding ways to incorporate this compositionality into the network structure is another direction that interests me. Similar to [Kotzias et al., 2015], I am currently working on a hierarchical LSTM to capture

the inter-dependence of sentences within a review. I am not only interested in how sentences can be composed into documents [Li et al., 2015], but also how other linguistic units such as different argumentative structures, e.g. the premise and hypothesis in entailment [Rocktäschel et al., 2015], the components of a summary, or the different markers of a conversation interact with each other. Ultimately, I am interested in developing models that compose knowledge and that are able to derive high-level concepts from basic assumptions, be it generalizing from examples, learning abstractions from individual characteristics, or learning high-level programs from sub-programs [Reed and de Freitas, 2015].

3 Dialogue and conversation

With the advent of the sequence-to-sequence framework, even classic NLP problems have been framed as sequence generation tasks, e.g. parsing [Vinyals et al., 2015] and NER [Gillick et al., 2015]. Generally, much of human learning is inherently sequential and is based on dialogue. To this effect, I am interested in finding better ways to model sequences that lend themselves to modeling dialogue and conversation. Representing dialogue in order to better capture the opponent’s intent and address her needs would be the primary goal; conversely, leveraging the dialogue-inherent feedback [Mikolov et al., 2015, Weston, 2016] in order to enable the agent to learn is another research direction that profoundly intrigues me.

References

- [Gillick et al., 2015] Gillick, D., Brunk, C., Vinyals, O., and Subramanya, A. (2015). Multilingual Language Processing From Bytes. *arXiv preprint arXiv:1512.00103*.
- [Gupta et al., 2015] Gupta, A., Boleda, G., Baroni, M., and Pad, S. (2015). Distributional vectors encode referential attributes. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2015)*, (September):12–21.
- [Iyyer et al., 2016] Iyyer, M., Guha, A., Chaturvedi, S., Boyd-graber, J., and Daum, H. (2016). Feuding Families and Former Friends : Unsupervised Learning for Dynamic Fictional Relationships. *NAACL*.
- [Kotzias et al., 2015] Kotzias, D., Denil, M., de Freitas, N., and Smyth, P. (2015). *From Group to Individual Labels using Deep Features*.

- [Le and Mikolov, 2014] Le, Q. and Mikolov, T. (2014). Distributed Representations of Sentences and Documents. *International Conference on Machine Learning - ICML 2014*, 32:1188–1196.
- [Li et al., 2016] Li, J., Galley, M., Brockett, C., Gao, J., and Dolan, B. (2016). A Persona-Based Neural Conversation Model. *arXiv preprint arXiv:1603.06155*.
- [Li et al., 2015] Li, J., Luong, M.-T., and Jurafsky, D. (2015). A Hierarchical Neural Autoencoder for Paragraphs and Documents. *ACL*, pages 1106–1115.
- [Mikolov et al., 2013] Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. *NIPS*, pages 1–9.
- [Mikolov et al., 2015] Mikolov, T., Joulin, A., and Baroni, M. (2015). A Roadmap towards Machine Intelligence. *arXiv preprint arXiv:1511.08130*.
- [Reed and de Freitas, 2015] Reed, S. and de Freitas, N. (2015). Neural Programmer-Interpreters. *arXiv preprint arXiv:1511.06279*, pages 1–12.
- [Rocktäschel et al., 2015] Rocktäschel, T., Grefenstette, E., Hermann, K. M., Kočiský, T., and Blunsom, P. (2015). Reasoning about Entailment with Neural Attention. *arXiv preprint arXiv:1509.06664*, pages 1–9.
- [Vinyals et al., 2015] Vinyals, O., Kaiser, L., Koo, T., Petrov, S., Sutskever, I., and Hinton, G. (2015). Grammar as a Foreign Language. *Advances in Neural Information Processing Systems*.
- [Weston, 2016] Weston, J. (2016). Dialog-based Language Learning. *arXiv preprint arXiv:1604.06045*.