Neural networks for modeling relations in text

Mots clés :
- Directeur de thèse : Ludovic Denoyer
- Co-encadrant(s) :
- Unité de recherche : Laboratoire d'informatique de Paris 6
- Ecole doctorale : École Doctorale Informatique, Télécommunications, Électronique de Paris
- Domaine scientifique principal: Divers

Résumé du projet de recherche (Langue 1)
Humans perceive the physical world according to spatial and temporal dimensions. They move in the physical world by encoding objects and their relations and expressing them using language. Contextual world knowledge is essential to understand linguistic expressions and relations referring to objects, actions and events. A fundamental question for understanding human language and its relation to the physical world is then how to build semantic representations for spatial and temporal language phenomena. A recent trend in language technology is the development of representation learning (Bengio 2013, Lecun 2015), a major topic in machine learning which has recently become a field by itself. Representation learning and its flagship, Deep Learning, have become in a few years key topics in natural language analysis. The purpose of the thesis is to examine the potential of this family of methods in the field of information extraction from text.

Résumé du projet de recherche (Langue 2)
Neural networks have recently witnessed some successes in this field with the development of word embeddings (Mikolov 2013, Socher 2014). Impressive results have been recently obtained for learning joint structured representations, allowing for example to provide the description of images via generated sentences (Vinyals 2014, Kiros 2014), nearly reaching human performance. Recurrent neural networks are state space dynamical models which offer a natural solution to the modeling of word sequences. A recent breakthrough in this domain is the development of gated recurrent networks which have opened the way for large scale applications of these techniques like handwriting or speech recognition (Graves 2013) or language translation (Cho 2014). Embeddings are tailored to capture semantics and encode semantic similarity. These models outperform other distributional models on different semantic similarity tasks (Baron 2014). The thesis objective is to develop new neural network models for analyzing and understanding human language in textual form. We will focus on the analysis and extraction of relations pertaining to spatial and temporal expressions in language. The thesis will be organized around three main tasks. * Improving semantic representation for words and expressions Classical representation learning has considered general multi-word embeddings. We will focus here on words and expressions specific of spatial and temporal information. For this the thesis will explore different techniques for combining different sources of information; raw text, knowledge bases, ontologies and resources like Wikipedia or DBpedia so as to enrich the target information terms and expressions. * Binary relation extraction on word sequences Relation extraction will be first explored for binary relations between textual elements. The thesis will develop sequence models based on recurrent neural networks with attention mechanisms (Mnih 2015) for analyzing sequences and extracting relations in a classification/ ranking framework. * Complex inferences for distant relations Relations are often between distant items in a text, or are non-explicit. Their detection requires performing complex inferences and make several steps of reasoning. This is a grand challenge in Artificial Intelligence. Recently, neural network architectures, called memory models have been proposed specifically with this type of task as targets (Kumar 2016, Sukhbaatar 2015). They are trained to learn sequence item representations together with some attention mechanism able to combine or select the representations relevant for a given task. These models have been used mainly on artificial tasks for now. The thesis will explore their use for learning to perform complex inferences on text and on the extraction of spatio-temporal relations. (Baron 2014) Baroni, M., Dinu, G., Kruszewski, G.: Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL). pp. 238–247 (2014) (Bengio 2013) Y. Bengio, A. Courville, P. Vincent, "Representation Learning: A Review and New Perspectives," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, pp. 1798-1828, Aug., 2013 (Kiros 2014) Kiros R., Salakhutdinov R. and Zemel R. S. Unifying visual-semantic embeddings with multimodal neural language models. In arXiv:1411.2539, 2014. (Kumar 2016) Kumar, A., Irsoy, O., Su, J., Bradbury, J., English, R., Pierce, B., … Socher, R. (n.d.). Ask Me Anything: Dynamic Memory Networks for Natural Language Processing., Arxiv. (Lecun 2015) LeCun, Y., Bengio, Y. and Hinton, G. E.,Deep Learning. Nature, Vol. 521, pp 436-444. (Mikolov 2013) Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J.: Distributed representations of words and phrases and their compositionality. In: Proceedings of the 27th Annual Conference on Advances in Neural Information Processing Systems (NIPS). pp. 3111–3119 (2013) (Mnih 2014) Mnih, V., Heess, N., Graves, A., & Kavukcuoglu, K. (2014). Recurrent Models of Visual Attention. In Advances in Neural Information Processing Systems 27 (Vol. 27, pp. 1–9). (Silberer 2014) Silberer, C., Lapata, M.: Learning grounded meaning representations with autoencoders. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL), pp. 721–732 (2014) (Socher 2014) Socher, R., Karpathy, A., Le, Q.V., Manning, C.D., Ng, A.Y.: Grounded compositional semantics for finding and describing images with sentences. Transactions of the Association for Computational Linguistics (TACL) 2, 207–218 (2014) (Sukhbaatar 2015) Sukhbaatar, Sainbayar; Szlam, Arthur; Weston, Jason; Fergus, Rob, End to End Memory Networks, NIPS 2015 (Vinyals 2014) Vinyals O., Toshev A., Bengio S., Erhan D.: Show and tell: a neural image caption generator, ArXiv 2014.