Does a dog desire cake? - Expanding Knowledge Base Assertions Through Deep Relationship Discovery

Anonymous Author(s)
Affiliation
Address
`email`

Abstract

Relationship discovery between two entities is a problem that has to be addressed when constructing a Knowledge Base (KB). A solution to this problem is important because the KB built from the discovered relations can play a key role in downstream tasks, such as analogical reasoning. An example of this kind of reasoning is whether a dog desires cake: a dog is an animal, cake is food, animals desire food therefore a dog desires cake. We constructed a system that that is trained on a commonsense KB and whose inputs are pairs of concepts and its outputs are the strength of commonsense assertions between the concepts. Our approach is unique because it can handle out of vocabulary entities and can generalize commonsense to out of knowledge concepts. We utilize the system to be able to infer the answer for out of knowledge assertions such as the aforementioned whether a dog desires cake.

1 Introduction

The problem that we set out to solve is the following. If we were given a set of possibly new entities, how could we extract how the entities relate to each other. Our approach, summarized here is the following. We first learn a FastText(1) word vector representation of entities. We then proceed to retrofit this representation with the information in a KB. In particular, we utilize the information found in ConceptNet which is a commonsense understanding of the world. Now, since our enriched word vectors only contain information found in the training text or explicitly stated in the KB, we proceed to generalize this knowledge. To accomplish this, we developed a CycleGAN(2) based system called RetroGAN that learns the mapping from word embeddings to retrofitted word embeddings. By learning this mapping, the system is learning to generalize the information in the knowledge graph by fusing it with the information present in the word embeddings. The interesting part about this is that as long as you can generate the word embedding, you will be able to generate its generalized retrofitted counterpart, and since we are using FastText(1), the generation of new out of vocabulary entries is relatively robust thanks to the sub-word information learned in the training of FastText. Additionally, since we have learned the mapping to the retrofitted counterpart, we are no longer limited to in-knowledge entities. An example of this is if our KB did not have the entity doggo. Doggo is internet slang for dog. With RetroGAN we can generate a retrofitted embedding for doggo that should have similar information to that of the dog embedding.

After we have this retrofitting mapping through RetroGAN we run into the problem that we need to be able to extract the learned knowledge to be able to build knowledge bases. To accomplish this, we built a system called Deep Relationship Discovery (DRD) whose inputs are pairs of learned-retrofitted word vectors, and its outputs are the strength of commonsense assertions between the two input concepts. Intuitively, DRD learns that semantically similar entities should have similar assertions. We developed a Graphical User Interface (GUI) with the intention of testing the inference from DRD.

Submitted to 33rd Conference on Neural Information Processing Systems (NeurIPS 2019). Do not distribute.

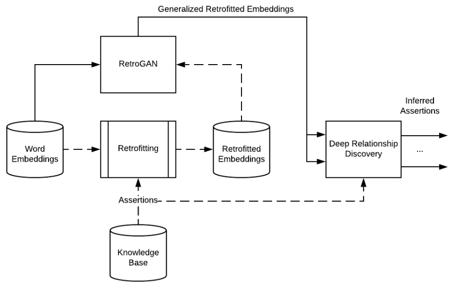

Figure 1: Complete System Architecture. The dashed lines represent procedures that are only necessary when the system is being trained.

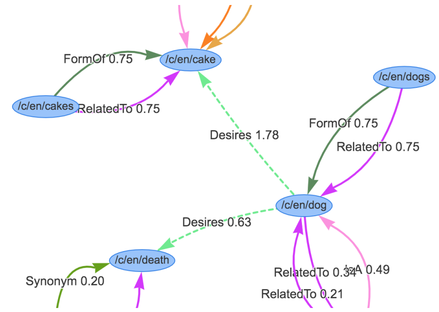

Figure 2: DRD Visualization tool. The blue ellipses are the concepts that we are visualizing, and the arrows are the the relationships between them. The solid arrows are existing relationships and the dotted arrows are inferred relationships.

In a simplistic example, in the GUI we can load the concept dog, the concept cake, and the concept death, and test the inference of whether a dog desires cake or if a dog desires death. This is shown in figure 2. The strength of the dog and cake assertion is shown to be 1.78 which is very different from if we tested whether a dog desires death which gives a strength of 0.63.

Putting it all together, with the combination of the RetroGAN and the Deep Relationship Discovery, if we generate a pair of word embeddings (on possibly new entities) and pass it through the RetroGAN system, we can get a an expanded commonsense-retrofitted representation of these pairs. We can then deconstruct this commonsense representation with the Deep Relationship Discovery. The result is how those two concepts relate within the context of common sense. If we iterate over all of the pairs of entities in a new topic, then our end result is a set of assertions that show how the entities in the new task relate from the perspective of commonsense.

2 Future Work

There are many areas that this work can be improved and continued. We intend to test our RetroGAN system by training it with Attract-Repel retrofitting strategies and evaluate it with downstream tasks such as lexical text simplification similar to what is done for AuxGAN (3) to understand better the effect of the CycleGAN architecture in learning the mapping. We intend to test our Deep Relationship Discovery system through human evaluation of previously unseen assertions. Additionally, we want to explore the optimization of the network configuration and to explore different ways to train the system by augmenting the data with some noise possibly to improve the generalization performance.

Looking at other areas, we want to leverage domain specific knowledge with general commonsense knowledge. To this end we are working on developing a transfer learning mechanism so that our system can adapt the commonsense understanding to some topic dependent knowledge. The reason for this is to leverage the connections and assertions that appear on a domain specific matter and combine it with the much broader commonsense information. If we were able to achieve this, we could build systems that can produce KBs that can be used for task-specific reasoning.

3 Conclusion

This work presents an expansion on work done to generalize retrofitting mappings though the use of a CycleGAN(2) system called RetroGAN. Additionally, we develop a novel way to discover commonsense-based assertions between entities, by training a Multi-Task Learning (MTL)(4) system on a subset of the assertions present in ConceptNet(5). We explored the combination of the RetroGAN system with the Deep Relationship Discovery one to be able to infer assertions from concepts that may or may not be in the vocabulary, and that may or may not be in the knowledge base. We utilize this system to be able to infer that a dog does indeed desire cake!

References

[1] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," *Transactions of the Association for Computational Linguistics*, vol. 5, pp. 135–146, 2017.

[2] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232, 2017.

[3] E. M. Ponti, I. Vulić, G. Glavaš, N. Mrkšić, and A. Korhonen, "Adversarial propagation and zero-shot cross-lingual transfer of word vector specialization," *arXiv preprint arXiv:1809.04163*, 2018.

[4] R. Collobert and J. Weston, "A unified architecture for natural language processing: Deep neural networks with multitask learning," in *Proceedings of the 25th international conference on Machine learning*, pp. 160–167, ACM, 2008.

[5] R. Speer, J. Chin, and C. Havasi, "Conceptnet 5.5: An open multilingual graph of general knowledge," in *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.