

Abstract

Word embeddings capture semantic relations as translation between vectors. This project proposes and evaluates a method that given a relation and an embedding it learns such transformation. 1-2 paragraphs summarizing the project topic, methods used and evaluation. Give a general idea, without details!

Introduction Explain the main objective of your project - what are you trying to do? how is it done now? How is this project different? Why is it an interesting problem?

Extensive work shows that word embeddings can capture semantic and syntactic relations mikolov2013efficient, mikolov2013linguistic, levy2014linguistic as vector offsets. Specifically, such embeddings are usually evaluated in similarity tasks (distance between related words) and analogy tasks (complete the sentence: “is to , as is to ”) where the type of relation is not explicitly provided. For a human to answer, the relation needs to be inferred and then applied it to the third argument levy2014linguistic. In nayak2015learning they propose a more specific task of predicting the hypernym of a given word. It can be seen as the task of completing the sentence “x is a y”.

We propose a variation of this task that is generic over the type of relation. For example: “a can ” or “a is cause of ”. This task is challenging in at least two different ways with respect to the analogy task: 1) word relations need not be functions: there might be two such instances of b that hold. e.g. a bee can sting and a bee can fly; and 2) polysemy in an argument e.g. a dog can bark, bark(make a loud noise) and bark(part of plant). This are pitfalls of word embeddings in general, but the analogy task alleviates it by possibly disambiguating with the third argument c .

A dataset extracted from ConceptNet is used to find answers to such task with vector algebra, akin to work in the analogy task. This can be described as a way to represent the symbolic knowledge of concepts and edges into the word embeddings, where there’s a correspondence between concept \sim embedding, and edge \sim affine transformation.

Novel Aspects This new formulation of the task as searching a single affine transformation for all triplets of a relation over an embedding space is, to the best of my knowledge, novel.

Both GloVe and word2vec approaches to this is through analogies that do not make the specific relation between terms explicit. This is a tradeoff between generality and interpretability.

Problem Definition

For concepts c_h, c_t with corresponding word representations $w_h, w_t \in V : R^d$ and a relationship between them r , we want to find a function $f_r : R^d \rightarrow R^d$ such that: $f_r(w_h) \approx w_t$