Compositional Distributional Semantic Models for Semantic Relatedness and Entailment

Sidharth Gupta
11714
Dept of CSE, IIT Kanpur

S Sai Krishna Prasad
11620
Dept. Of CSE, IIT kanpur

Guide: Prof. Amitabha Mukerjee (CS365 - Spring 2014)

Introduction

Distributed semantic models approximate the meanings of words by studying the distribution of the word across different contexts in the given training data. This distribution characterises lexical semantics (Distributional Hypothesis) and is specified for every word in the form of a vector in high dimensional space. Compositional distributional semantic models seek to extend this approach to characterise the semantics of phrases or sentences, which simpler DSMs fail to model as they ignore grammatical structure and logical words, due to their non compositional nature.

The aim of our project is to use compositional DSMs to tackle the tasks of semantic relatedness (estimating how closely related linguistic items are in the semantic sense) and textual entailment (determining directional relationships between text fragments such as entailment or contradiction) on the SICK data set.
This is the first task in SemEval-2014 (http://goo.gl/2TL5ll)

MOTIVATION

Determining semantic relatedness in sentences has been one of the most researched topic in the field of Natural Language Processing (NLP). Human beings are able to understand the meanings of words which we have never encountered whereas machines have failed to do so. Search engines like Google and Yahoo fail when they have to find the semantic meaning of the phrases that we write. By establishing a process of understanding the semantics of the sentences these search engines will be able to perform applications like document search and query resolution.

RELATED WORK

In Grefenstette's and Sadrzadeh's paper[2], the authors implement a compositional DSM, training it over the entire BNC. The evaluation is based on the word disambiguation task developed by Mitchell and Lapata[3], and the results obtained match or better those of other competitors. They take an unsupervised learning approach to learn the matrices corresponding to relational words (suitably identified) as well as the distribution vectors corresponding to other words. Relational words are modeled as matrices to allow them to act on the vectors corresponding to the semantics of other words. This is needed while composing the semantics of the sentence as a whole, which is a function (linear map) of the Kronecker product of the word vectors.

"Semantic Compositionality through Recursive Matrix-Vector Spaces" is one of the works which captures the compositional meaning of longer sentences, a thing which other papers lack. The authors introduced a Recursive Neural Network (RNN) model which learns compositional meanings for longer phrases and sentences. Here the author represented each word using a vector and a matrix. The vector contains the meaning of the word whereas the matrix tries to explain how the word tries to modify the meaning of the words associated with it. The author created a new matrix and the corresponding vector by taking two words or phrases at a time and then recursively running them for the whole sentence.

DATASET

We will be using the SICK data set which consists of 10,000 English sentence pairs, divided into training and test sets of 5000 pairs each.Each pair in the training set is labeled for relatedness in meaning (on a scale of 1 to 5) as well as the entailment relation between the two sentences. The entailment relationship is defined as follows:- Given two sentences A and B we have to deduce if A implies B, B contradicts A or B and A are neutral with respect to each other. The dataset can be found at the following link:- http://goo.gl/YWBuXk

REFERENCES

1. Baroni, Marco and Roberto Zamparelli, 2010. Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space. In *Proceedings of EMNLP.* Boston, MA.
2. Grefenstette, Edward and Mehrnoosh Sadrzadeh, 2011. Experimental support for a categorical compositional distributional model of meaning. In *Proceedings of EMNLP*. Edinburgh, UK. (Recommended by Prof. Mukerjee)
3. Mitchell, Jeff and Mirella Lapata, 2008. Vector-based models of semantic composition. In Proceedings of ACL. Columbus, OH.
4. Mitchell, Jeff and Mirella Lapata, 2010. Composition in distributional models of semantics. Cognitive Science, 34(8): 1388–1429.
5. Socher, Richard, Brody Huval, Christopher Manning, and Andrew Ng, 2012. Semantic compositionality through recursive matrix-vector spaces. In Proceedings of EMNLP. Jeju Island, Korea. (Recommended by Prof. Mukerjee)