Word Embeddings

THE
UNIVERSITY
OF UTAH

Overview

- Representing meaning
- Word embeddings: Early work
- Word embeddings via language models
- Word2vec and Glove
- Evaluating embeddings
- Design choices and open questions

Overview

- Representing meaning
- Word embeddings: Early work
- Word embeddings via language models
- Word2vec and Glove
- Evaluating embeddings
- Design choices and open questions

The evaluation problem

- Suppose we have a way to convert words to vectors
  - Pick your favorite method

- The (sometimes unstated) implication here is that these vectors represent the meaning of words

- How can we verify this claim?

Thoughts?

Using word embeddings

Once we have word embeddings, what can we do with them? Several possibilities:

1. Measure word similarities and distances

   Eg: Cosine similarity of two words A and B = $\frac{\mathbf{a}^T\mathbf{b}}{||\mathbf{a}||\ ||\mathbf{b}||}$

   Other similarity functions are possible

2. Use this to find similar words or most dissimilar words

   Eg: Find the odd word among the following: `cat`, `tiger`, `dog`, `table`

Using word embeddings

Once we have word embeddings, what can we do with them? Several possibilities:

3. Document or short snippet similarities

   Question: If we have word vectors, how do we represent documents in the same vector space?

   Several answers. Most common: average or add the word embeddings

   Gives natural definitions for document similarities

Two broad families of evaluations

1. Intrinsic evaluation: Evaluate the representation directly without training another model

2. Extrinsic evaluation: Evaluate the impact of the representation on another task

Two broad families of evaluations

1. Intrinsic evaluation: Evaluate the representation directly without training another model
   - Typically simple tasks where success or failure is (almost) entirely a function of the representation
   - Easy to compute, but doesn't say much about the embeddings as features

2. Extrinsic evaluation: Evaluate the impact of the representation on another task

Two broad families of evaluations

1. Intrinsic evaluation: Evaluate the representation directly without training another model
   - Typically simple tasks where success or failure is (almost) entirely a function of the representation
   - Easy to compute, but doesn't say much about the embeddings as features

2. Extrinsic evaluation: Evaluate the impact of the representation on another task
   - Typically, a neural network
   - Can be more practically useful, but slow and depends on the quality of the model for the task being tested

Word Analogies

Given an incomplete analogy of the form

a : b :: c : ?

Find the word that best answers fits

The famous example:

King : Queen :: Man : ?

Word Analogies

Given word embeddings, one way to answer the question "a : b :: c : ?" is

$$\operatorname{argmax}_i \frac{(x_a - x_b + x_c)^T x_i}{||x_a - x_b + x_c||}$$

That is, if the answer is the word d, then we have

$$x_a - x_b \approx x_c - x_d$$

Word Analogies

Given word embeddings, one way to answer the question "a : b :: c : ?" is

$$\mathrm{argmax}_i \; \frac{(x_a - x_b + x_c)^T x_i}{||x_a - x_b + x_c||}$$

That is, if the answer is the word d, then we have

$$x_a - x_b \approx x_c - x_d$$

Not the only way to answer the question. Instead of this additive method, we could do something multiplicative

Word analogies data sets

Several standard datasets exist for word analogies

- Some capture syntactic patterns
    - give : giving :: take : ?
- Some capture semantic patterns
    - queen: king :: tigress : ?
- Some require world knowledge
    - Utah : Salt Lake City :: Iowa : ?

General trends

- More data helps with analogy evaluations

- Very low and very high dimensional vectors seem to work
  - Need a sweet spot for best results

Word similarity evaluation

- Another intrinsic evaluation

- Pairs of words are hand-annotated with similarity scores

- The goal of the embeddings is to produce these scores
  - Or perhaps more reasonably, similar clusterings or rankings as the scores

- Standard software libraries exist for evaluating embeddings in this fashion