Lecture 11: Vector Semantics

Generated by LectureMate

February 21, 2025

1 Introduction

In this lecture, we explore vector semantics, focusing on the meanings of words and how they can be inferred from context. This builds upon previous discussions about mapping words to objects, particularly in the context of children's language acquisition.

2 Recap of Previous Concepts

We previously discussed how children learn new nouns by associating them with objects in their immediate context. This process involves several biases or priors, such as:

- Whole object bias
- Taxonomic bias
- Mutual exclusivity assumption

These biases help make the learning problem tractable by providing prior information that guides the mapping of words to meanings.

3 Bayesian Model of Number Learning

We examined a Bayesian model of number learning in children aged 20 to 40 months. The model involves:

- Input pairs that associate words with sets (e.g., "three" with three cookies).
- Hypothesis testing to determine the mapping procedure from sets of objects to number words.

The model computes the probability of hypotheses given observed data, using Bayes' theorem:

$$P(H|W,S) \propto P(W|S,H) \cdot P(H)$$

where H represents hypotheses, W the words, and S the sets.

4 Learning Mechanism

The likelihood term is defined based on the ability to generate correct number words, while the prior is influenced by the complexity of the hypothesis. Simpler hypotheses are preferred, aligning with the minimum description length principle.

5 Word Learning and Context

We transitioned to discussing word learning beyond nouns, including abstract nouns and function words. The key idea is that the meaning of a word can often be inferred from its context. This leads to the concept of word vectors or embeddings.

6 Word Vectors and Context

Word vectors are numerical representations of words based on their context in large text corpora. The process involves:

- Defining a context window around target words.
- Counting co-occurrences of target words with context words to form context vectors.

These vectors allow for various applications, including:

- Arithmetic operations on vectors
- Clustering similar words
- Input for neural networks

7 Cosine Similarity and Distance Measures

To determine the similarity between word vectors, we can use:

- Euclidean distance
- Cosine similarity

Cosine similarity is particularly useful for measuring how similar two vectors are, regardless of their magnitude.

8 Learning Word Embeddings with Neural Networks

We discussed the Word2Vec model, which uses a neural network to learn word embeddings. The architecture includes:

- Input layer representing context words
- Hidden layer that produces the word embedding
- Output layer predicting the target word

This model leverages self-training, where the correct answers are derived from the training data itself.

9 Applications of Word Embeddings

Word embeddings enable various linguistic tasks, such as:

- Finding synonyms and antonyms
- Performing arithmetic operations on words (e.g., ”king” - ”man” + ”woman” = ”queen”)
- Identifying relationships between words based on vector proximity

10 Conclusion

In summary, we explored how vector semantics provides a framework for understanding word meanings through context. The use of Bayesian models and neural networks allows for effective learning of word representations, which can be applied in various linguistic and computational tasks.