

Part-of-Speech Tagging with Word Embeddings

CS 9875 Final Project

Paul Moore Juhani Dickinson

December 3, 2025

Outline

Part-of-Speech Tagging

Word Embedding Models

Evaluation and Theoretical Analysis

Part-of-Speech Tagging

Parts of Speech

- ▶ Each word in a sentence carries out a syntactic role: denote an object, denote an action, modify an object, etc.
- ▶ This role is mapped to the word using a part-of-speech (PoS) tag

Spelled the Same, Used Differently: Syntactic and Semantic Differences

Words can be different parts of speech (do different things) depending on where they are in the sentence and what is around them

1. Noun-Verb: “spot”

- ▶ “Your nose has a spot on it.” (N)
- ▶ “Can you spot him?” (V)

2. Noun-Adjective: “red”

- ▶ “The red I like is that one.” (N)
- ▶ “The red car has arrived.” (A)

Tagging

Given a sentence, assign a PoS tag to each word in the sentence.

Can you spot the large spot on your nose ?

Tagging

Given a sentence, assign a PoS tag to each word in the sentence.

Can	you	spot	the	large	spot	on	your	nose	?
VB	D	VB	D	Adj	N	PP	D	N	.

Tagging

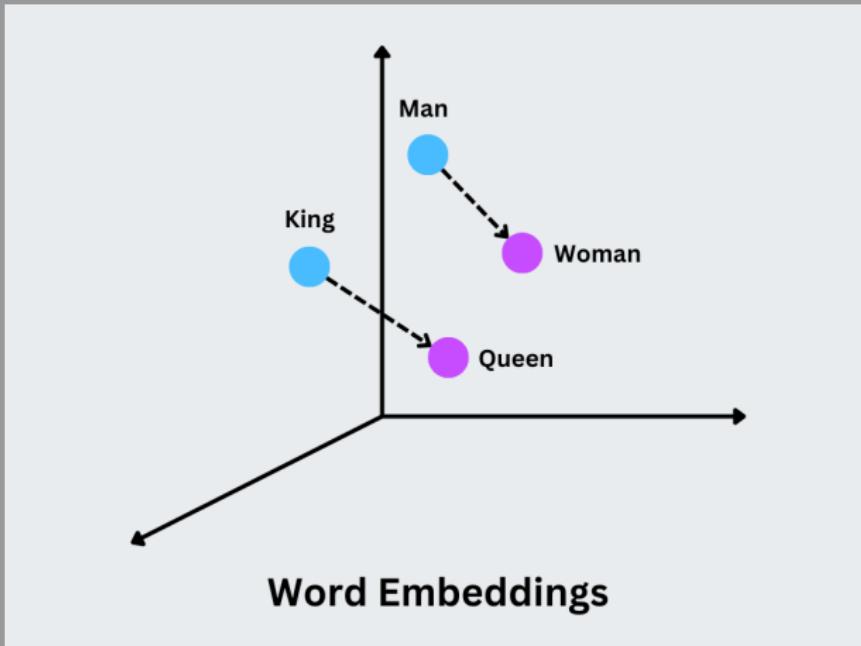
Given a sentence, assign a PoS tag to each word in the sentence.

Can	you	<u>spot</u>	the	large	<u>spot</u>	on	your	nose	?
VB	D	VB	D	Adj	N	PP	D	N	.

What makes PoS tagging challenging: context is very important.

Word Embedding Models

What are word embeddings?

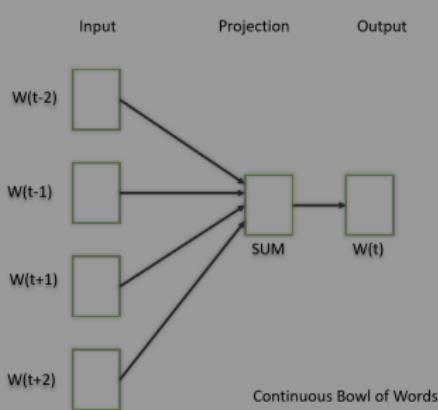


https://assets.zilliz.com/Figure_Word_EMBEDDINGS_b021a5a759.png

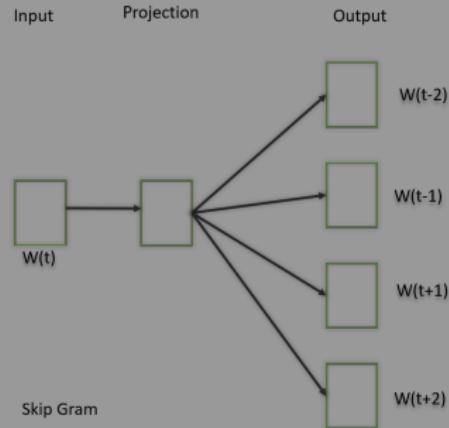
Static Word Embeddings

What are static word embedding models?

Word2Vec



CBOW



Skip Gram

Word2Vec Training Procedures
<https://www.geeksforgeeks.org/nlp/word-embeddings-in-nlp/>

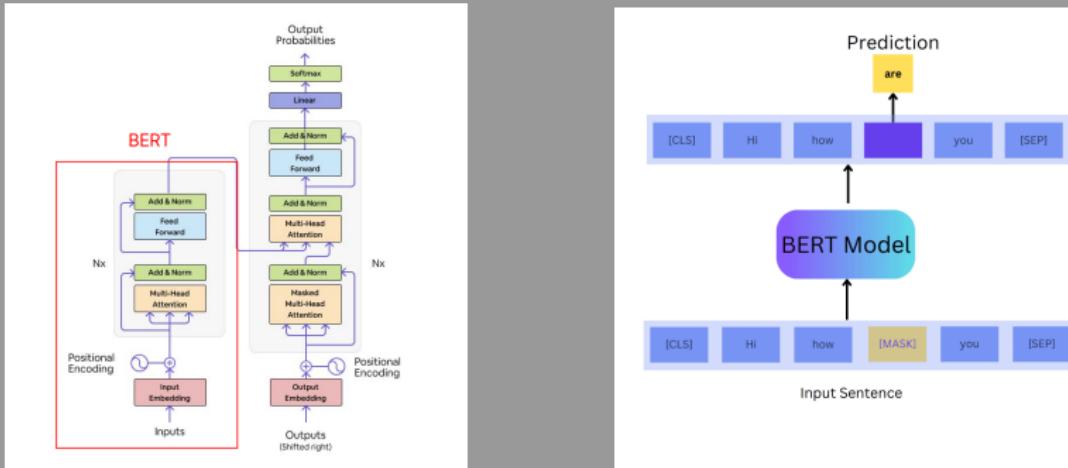
Word2Vec

$$\begin{bmatrix} W_{00} & W_{01} & W_{02} & \dots \\ W_{10} & W_{11} & W_{12} & \dots \\ W_{20} & W_{21} & W_{22} & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

Contextual Word Embeddings

What are contextual word embeddings?

BERT



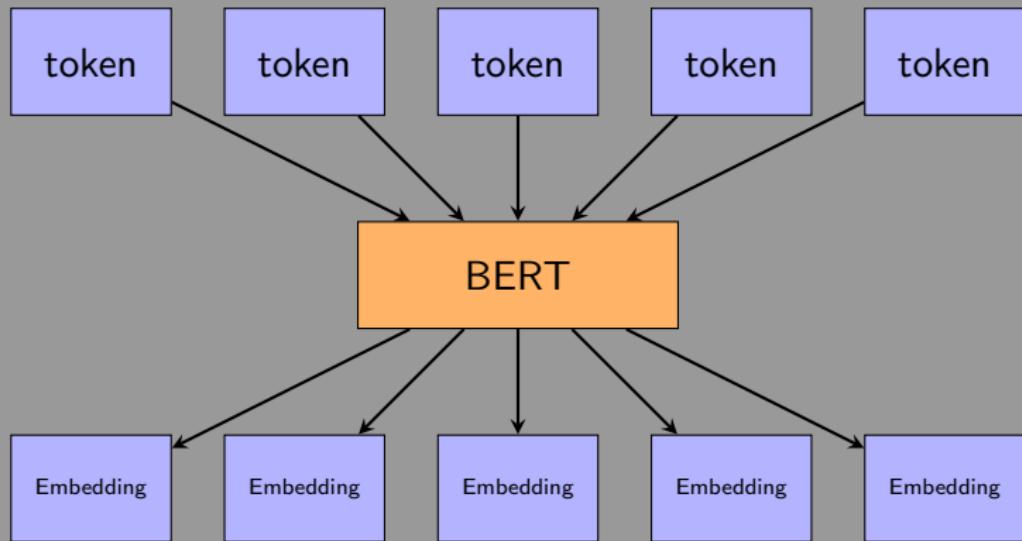
Transformer architecture

MLM

Transformer architecture: <https://deeplobe.ai/wp-content/uploads/2021/04/1.jpg>

MLM: <https://learnopencv.com/wp-content/uploads/2023/10/bert-masked-language-modeling-1.png> Devlin, Jacob; Chang, Ming-Wei; Lee, Kenton; Toutanova, Kristina, 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

BERT



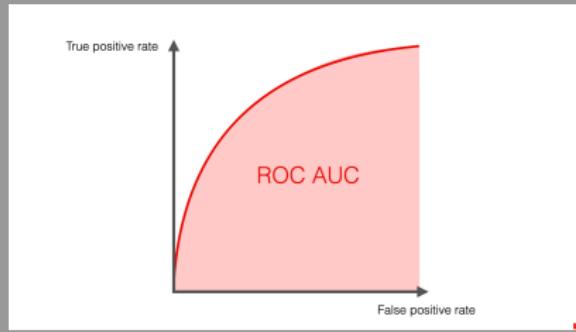
How do we use these embeddings?

How do we use these embeddings?
As input for a downstream model: SVM, Boosting, CNN, etc.

Evaluation and Theoretical Analysis

Evaluation Metrics

We will evaluate each of our models using F1 score and AUC-ROC adapted for this multi-class classification problem.



AUC-ROC

https://uploads-ssl.webflow.com/6266b596eef18c1931f938f9/64760748f6cfb67f889321ad_classification_metrics_016-min.png

Dataset Used

Brown Corpus:

- ▶ 1 million tagged words, validated by humans
- ▶ American English
- ▶ Provided by the Natural Language ToolKit (NLTK)

Theoretical Analysis

We will conduct a theoretical analysis on the word embedding models.

Anticipated Results

We anticipate achieving near perfect scores in both evaluation metrics using embeddings from contextual models like BERT and a simple classification model like an SVM, outperforming static embedding models like Word2Vec.