

Hide menu

Lecture: Word Embeddings

- ✓ **Video:** Week Introduction
1 min
- ✓ **Video:** Overview
2 min
- ✓ **Reading:** Overview
- ✓ **Video:** Basic Word Representations
3 min
- ✓ **Reading:** Basic Word Representations
5 min
- ▶ **Video:** Word Embeddings
3 min
- ✓ **Reading:** Word Embeddings
4 min
- ▶ **Video:** How to Create Word Embeddings
3 min
- ✓ **Reading:** How to Create Word Embeddings?
4 min
- ▶ **Video:** Word Embedding Methods
3 min
- ✓ **Reading:** Word Embedding Methods
4 min
- ▶ **Video:** Continuous Bag-of-Words Model
4 min

Week 4 Training a CBOW Model: Cost Function

Previous Next

Training a CBOW Model: Cost Function

The cost function for the CBOW model is a cross-entropy loss defined as:

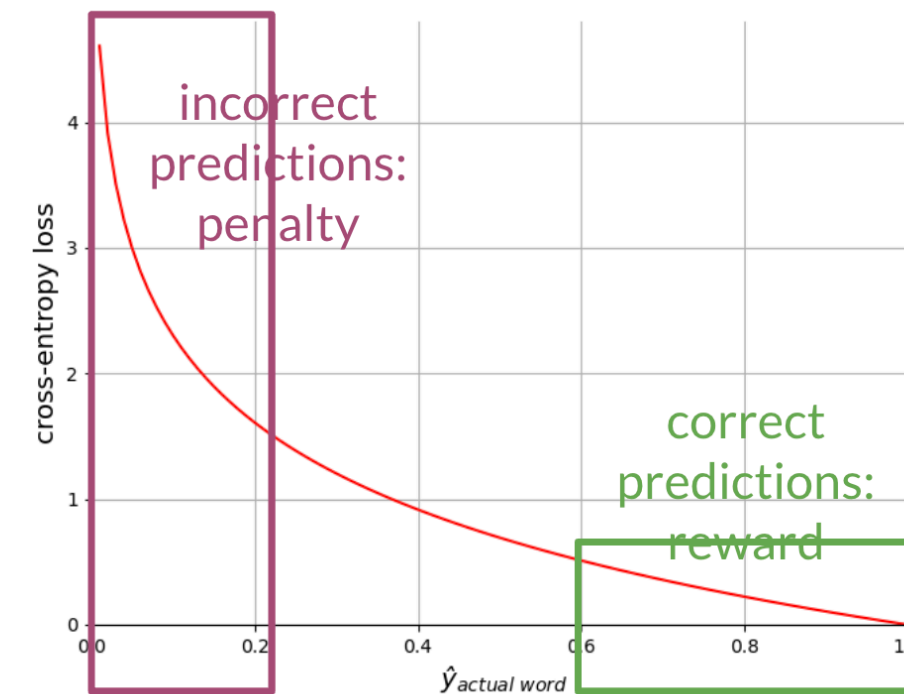
$$J = - \sum_{k=1}^V y_k \log \hat{y}_k$$

Here is an example where you use the equation above.

$$J = -\log \hat{y}_{\text{actual word}}$$

y		\hat{y}
0	am	0.96
0	because	0.01
1	happy	0.01
0	I	0.01
0	learning	0.01

→ J = 4.61



Why is the cost 4.61 in the example above?

