Hide menu

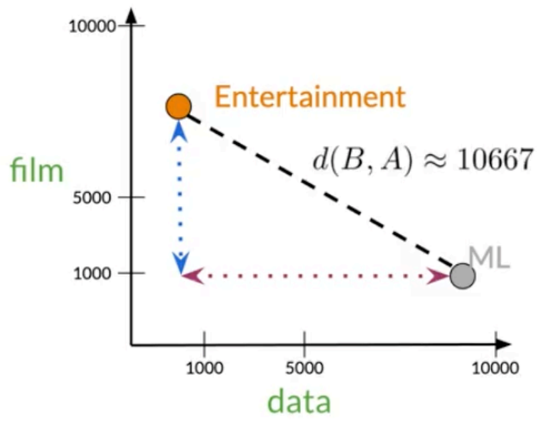
Lecture: Vector Space Models

- **Video:** Week Introduction
  47 sec
- **Video:** Vector Space Models
  2 min
- **Reading:** Vector Space Models
  10 min
- **Video:** Word by Word and Word by Doc.
  4 min
- **Reading:** Word by Word and Word by Doc.
  10 min
- **Lab:** Linear algebra in Python with Numpy
  1h
- **Video:** Euclidean Distance
  3 min
- **Reading:** Euclidian Distance
  10 min
- **Video:** Cosine Similarity: Intuition
  2 min
- **Reading:** Cosine Similarity: Intuition
  10 min
- **Video:** Cosine Similarity
  3 min
- **Reading:** Cosine Similarity
  10 min
- **Video:** Manipulating Words in Vector Spaces
  3 min
- **Reading:** Manipulating Words in Vector Spaces
  10 min
- **Lab:** Manipulating word embeddings
  1h
- **Video:** Visualization and PCA
  3 min

Week 3 > Euclidian Distance

Previous Next

Euclidean distance



Corpus A: (500,7000)

Corpus B: (9320,1000)

$$d(B, A) = \sqrt{(B_1 - A_1)^2 + (B_2 - A_2)^2}$$

$$c^2 = a^2 + b^2$$

$$d(B, A) = \sqrt{(8820)^2 + (-6000)^2}$$

You can generalize finding the distance between the two points (A, B) to the distance between an n dimensional vector as follows:

$$d(\vec{v}, \vec{w}) = \sqrt{\sum_{i=1}^{n} (v_i - w_i)^2}$$

Here is an example where I calculate the distance between 2 vectors ($n = 3$).

	data	$\vec{w}$ boba	$\vec{v}$ ice-cream
AI	6	0	1
drinks	0	4	6
food	0	6	8

$$= \sqrt{(1-0)^2 + (6-4)^2 + (8-6)^2}$$

$$= \sqrt{1+4+4} = \sqrt{9} = 3$$