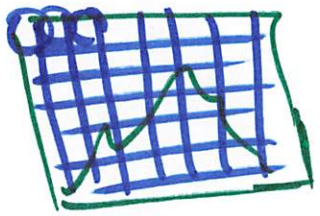
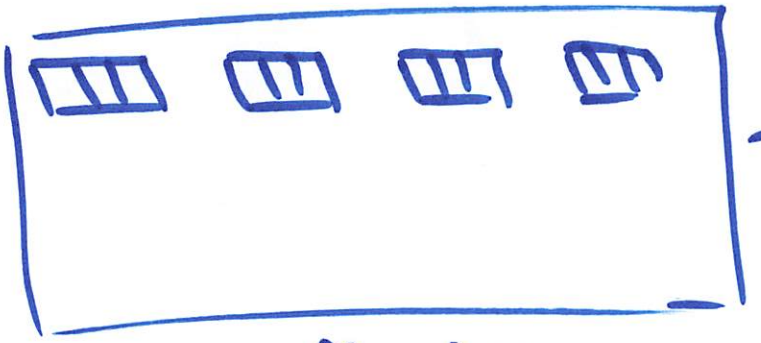


Img



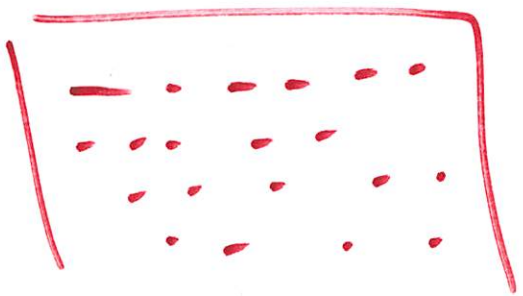
Numbers

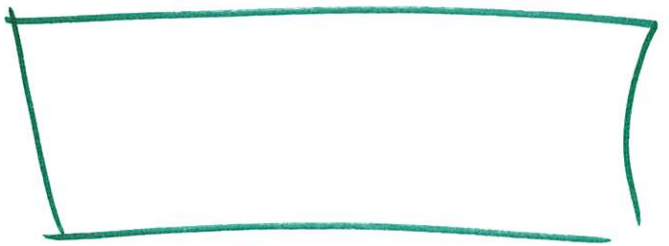
Math

Crop

Bright

rotating





Numbers

Math

translate

Sentim

...

One hot encoding

Euclidean

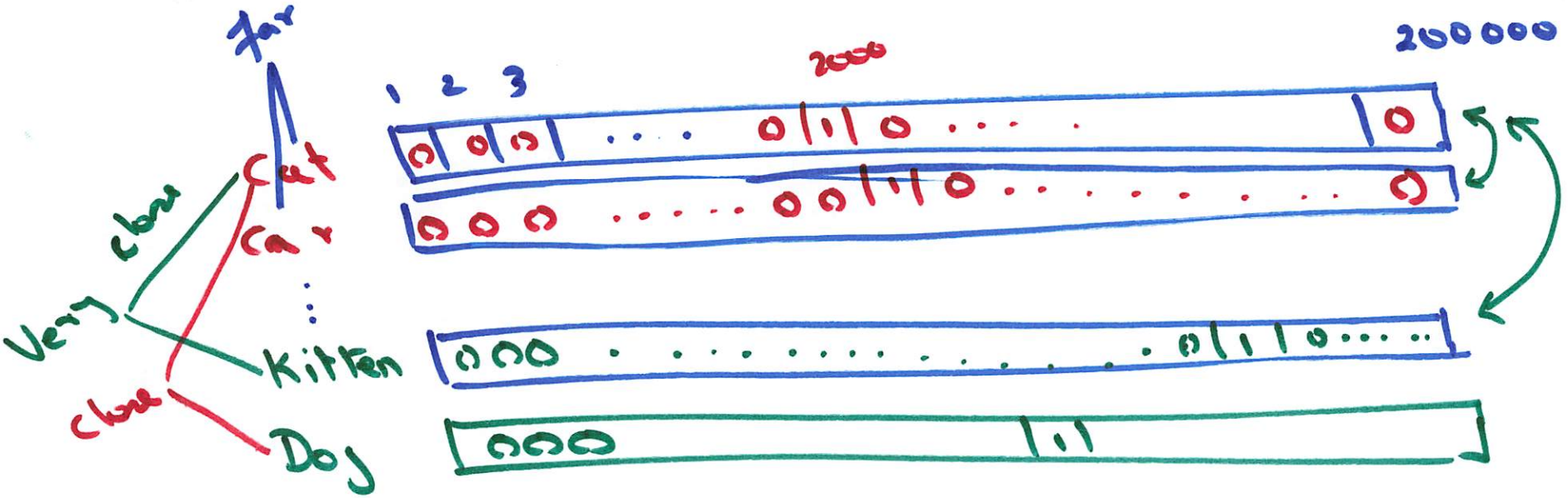
$\sqrt{\Delta x^2 + \Delta y^2}$

Δy

Δx

Cos Similarity

200,000 words



Word Embeddings

- Word Embeddings: An NLP technique of representing words as vectors
- Several methods exist to do this kind of translation. Popular methods include
  - One Hot encoding
  - Dimension reduction techniques, like SVD
    - word2Vec (Neural networks)
    - GloVe

Words → Vectors

7 million article
5 billion words
250 million sent!

- I went to the cinema on Sunday
- I went to the beach on Monday

- My favorite pet is a cat
- I like walking my pet dog

- I bought a red Honda
- I bought a red car
- I bought a blue Toyota

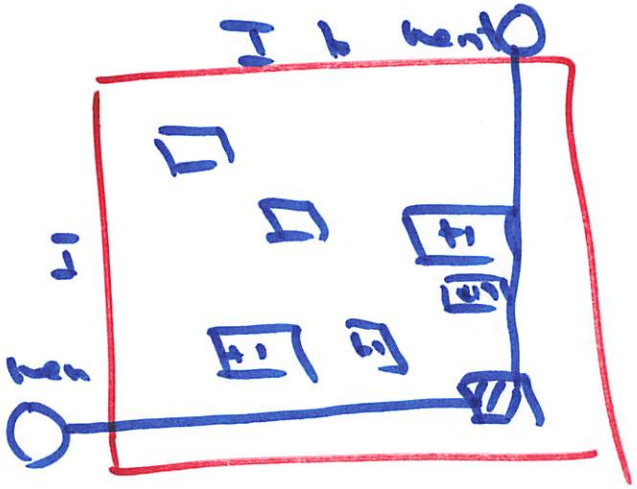
Dim Reduction

Co Occurance Matrix

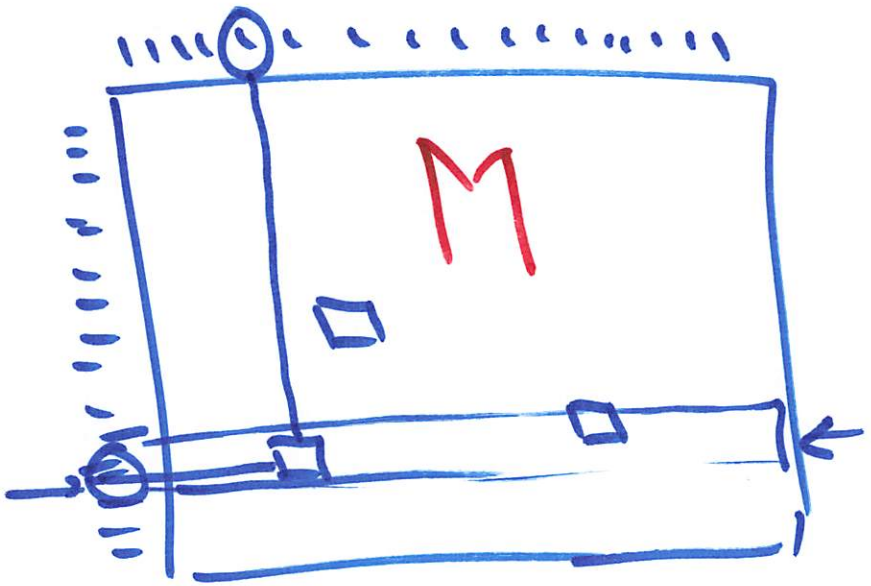
7 million

5 billion words

250 million sent.

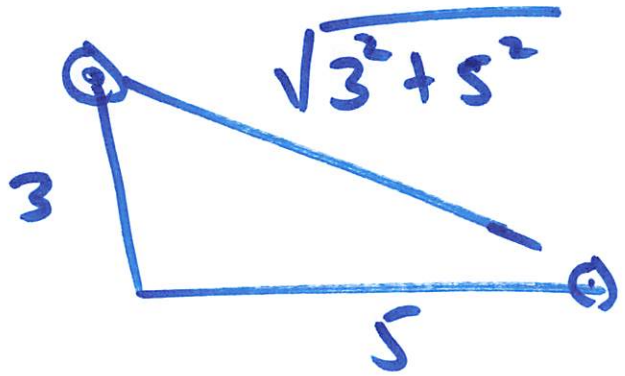


I went to the beach on Sunday.

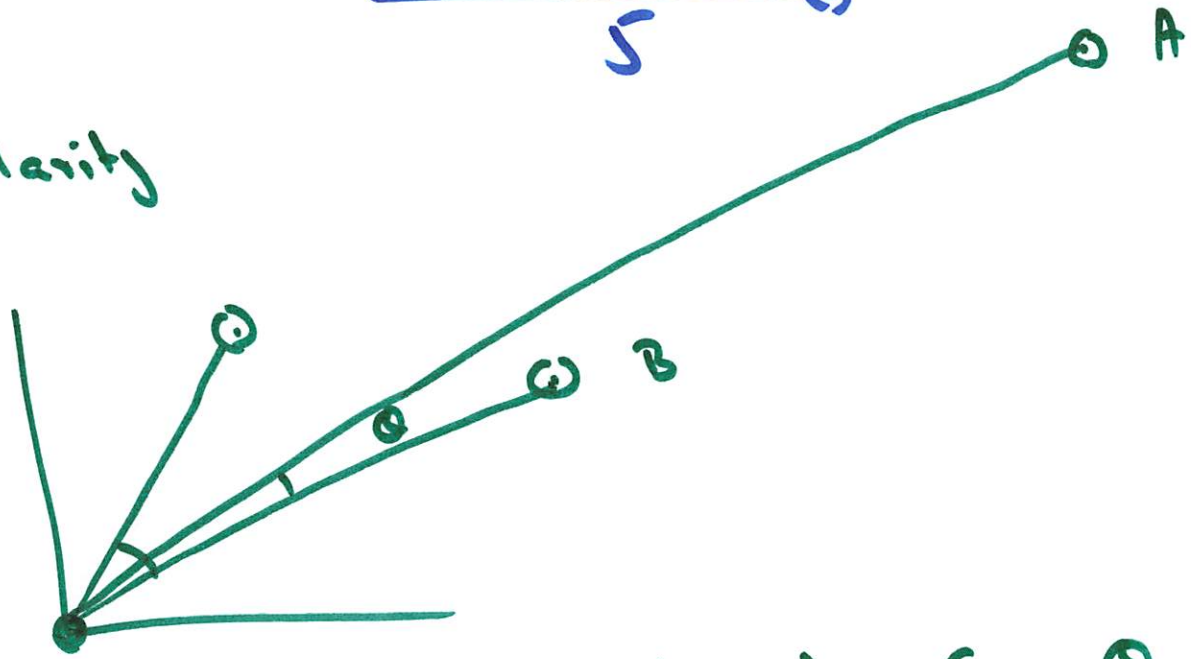


200000 x 200000

Euclidian



Cosine Similarity



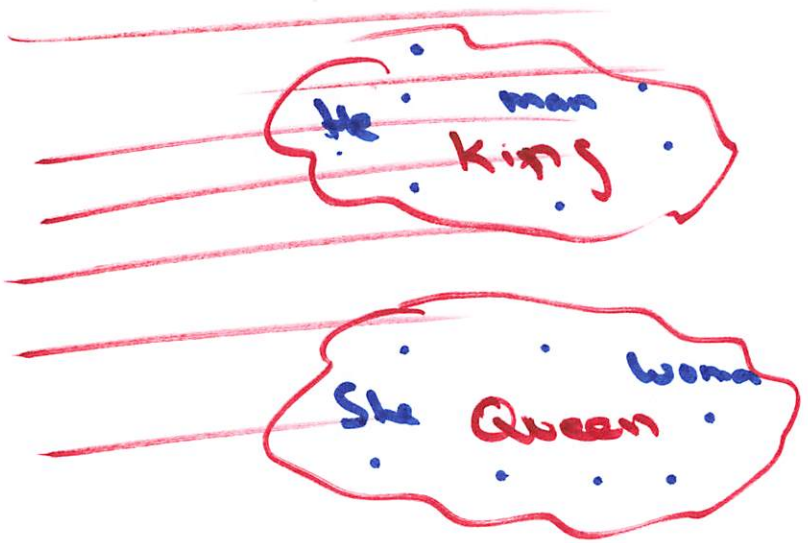
Similarity (A, B) = Cos θ

Cos θ

~~King : Man~~

Man : King as Woman : ? → Queen

King − Man + Woman = Queen



UK : London as US : ~~NY~~