

(25)

Embedding

In natural language processing, word embedding is a term used for the representation of words for text analysis, typically in the form of a real-valued vector that encodes the meaning of the word such that the words that are closer in the vector space are expected to be similar in meaning.

sentiment

review 1 → 20
 review 2 → 2000

[- - -]

sparse repres

0

pudding

sparse

20 → 1980 0

2000 num
 1980 → 0

this is nice

[0.7 0.1 0.3]
 non-zero

dense → dimensions
 ↓ 82 d'ims

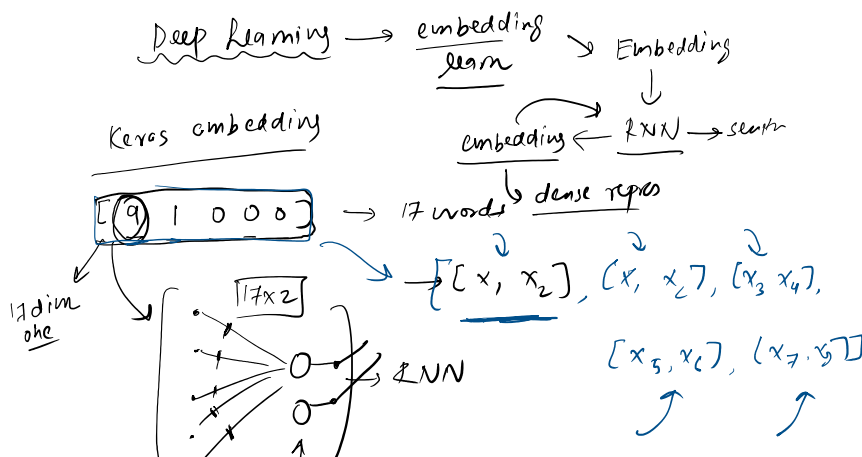
semantic meaning

word → context

WordVec
 Glove

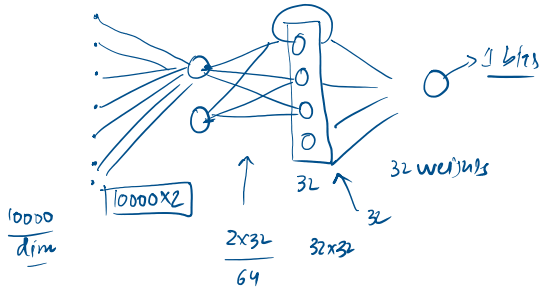
Silly for seq length like when padded then so many 0's , lead to sparse , check review wala example. Also these non-zero values , denote some semantic meanings.

For example : (if 3 values ,
then one may denote
grammar , punctuation etc ,
just as a example)





7 nodes



Step 1: Input X \Rightarrow (2,3) \Rightarrow batch_size=2, seq_len=3
 Step 2: Weight Matrix \Rightarrow (8,4) \Rightarrow embedding_dim=4
 Step 3: Output \Rightarrow (2,3,4) \Rightarrow embeddings for each token

Actually its like loopup which access index ,
 there is no such one hot encoding happens inside.

population in 9th year $\leftarrow x$

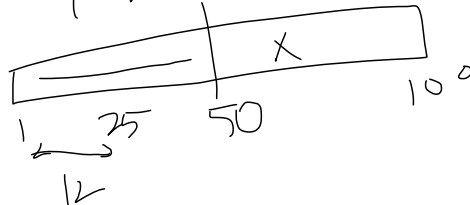
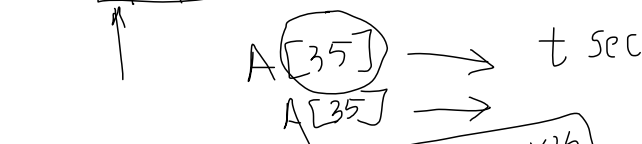
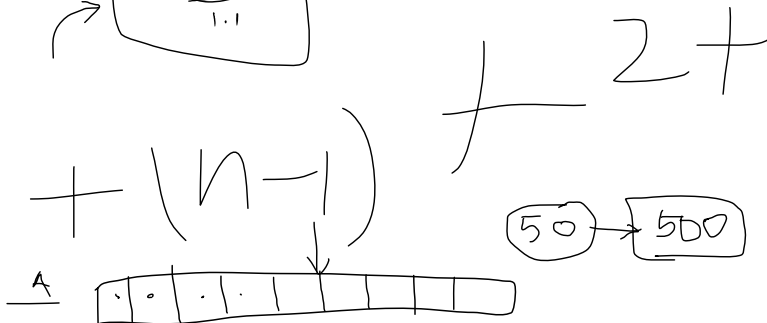
$$\frac{x + 10\% \text{ of } x}{x + 0.1x} = 10000$$

$t(n-1)$

$$\frac{1.1x}{1.1} = 10000$$

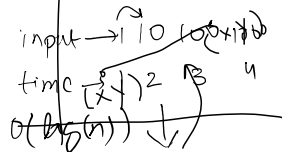
$$x = \frac{10000}{1.1}$$

$$\frac{x-1}{x} + \frac{1}{2} \left(\frac{x-1}{x} \right)^2 + \frac{1}{3} \left(\frac{x-1}{x} \right)^3 + \frac{1}{4} \left(\frac{x-1}{x} \right)^4 + \dots$$



$O(n)$

$O(n^2) \rightarrow$ nested loops



Binary search

