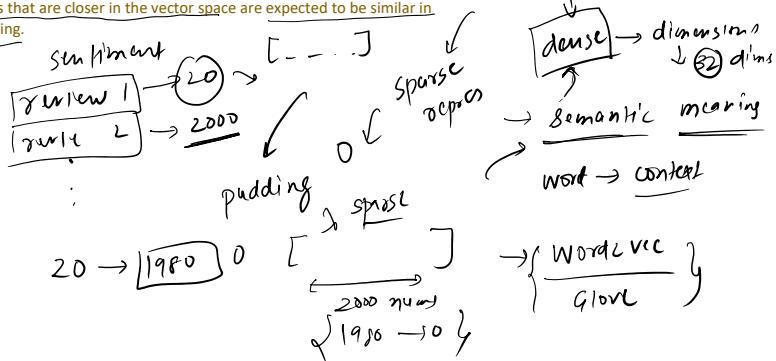


In natural language processing, word embedding is a term used for the representation of words for text analysis, typically in the form of a real-valued vector that encodes the meaning of the word such that the words that are closer in the vector space are expected to be similar in meaning.

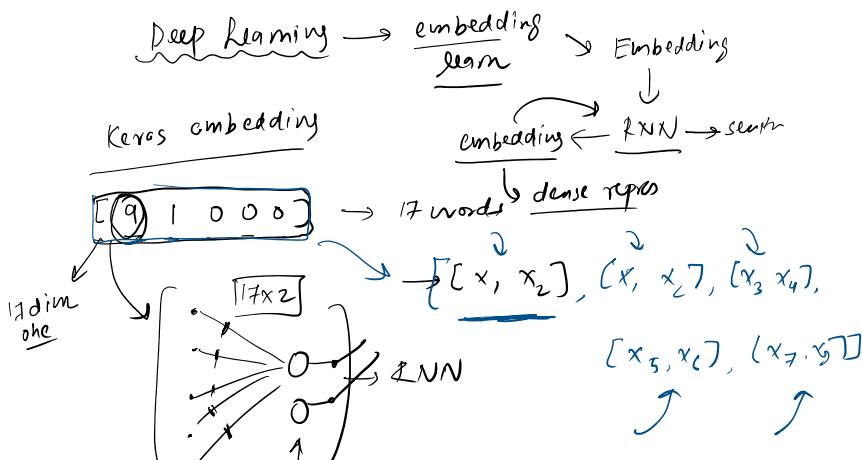


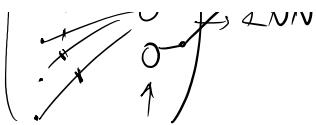
Let say vocab size is 1000, then a word is represented as [0, 1, ..., 0] (sparse), but in embedding it is like [0, 0.7, 0.1, 0.3] (dense).

Silly for seq length like when padded then so many 0's, lead to sparse, check review wala example.

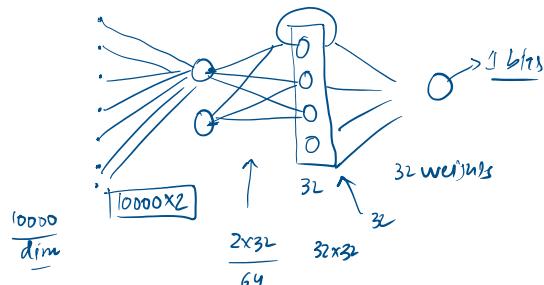
Also these non-zero values, denote some semantic meanings.

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





17 nodes



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

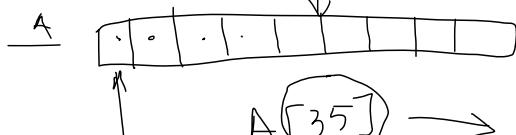
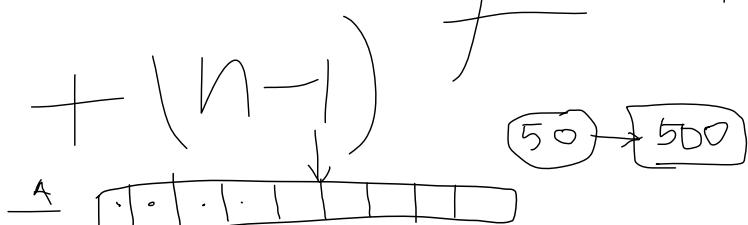
Actually its like loopup which access index ,
 there is no such one hot encoding happens inside.

population in 9th year $\in X$

$$\begin{aligned} X + \frac{10\% \text{ of } X}{X + 0.1X} &= 10000 \\ \Rightarrow X + 0.1X &= 10000 \\ \frac{1.1X}{1.1} &= 10000 \\ X &= 10000 \end{aligned} \quad t(n-1)$$

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

2+



A [35] \rightarrow t sec

A [35] \rightarrow

35 \rightarrow 1 x 4 x 35

1 25 50 100

12

$O(n)$

$O(n^2) \rightarrow \text{nested loops}$

Input \rightarrow 10 loops \times 10 loops
 $t(n) = (2)(n)^2 \approx n^2$
 $O(\frac{t(n)}{10})$

$\sqrt{(x_1)}$

Binary Search

1 — 100
 1 — 6