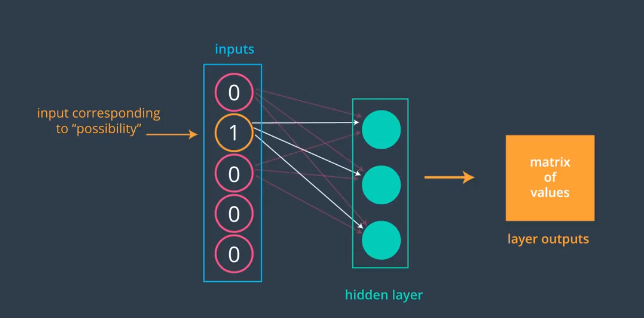


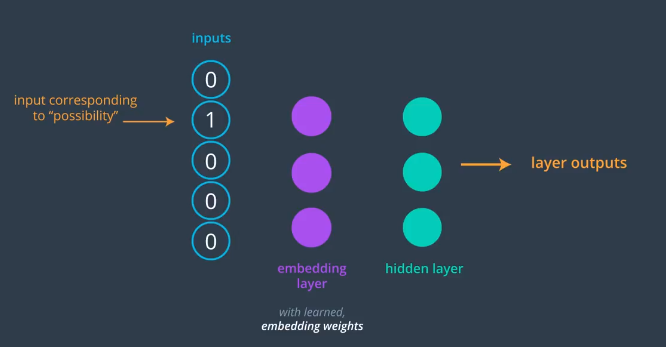
If text has bias like more man than woman, it will make error

Text to Data:

One hot encoding



Note: Since 1 hot encoding is expensive in text environment, we can use embedding layer before the hidden layer, the witghts of embedding layers



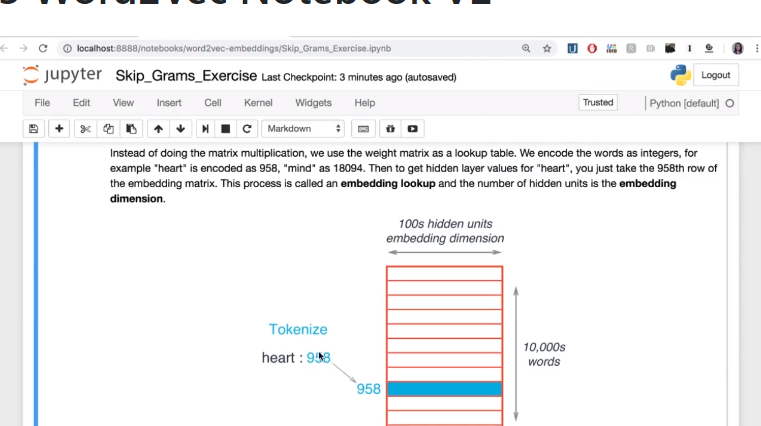
Each word as unique integer

Heart as 958, We just get 958th row as belove:



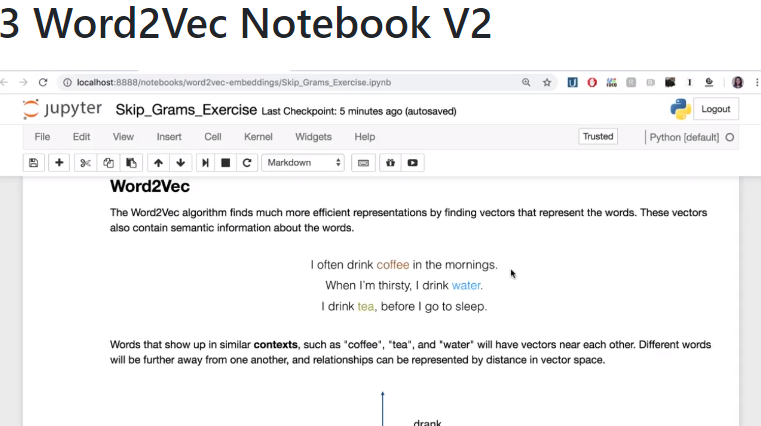
This number of hitting units.

In the case of large vocab, we remove the one hot encoding

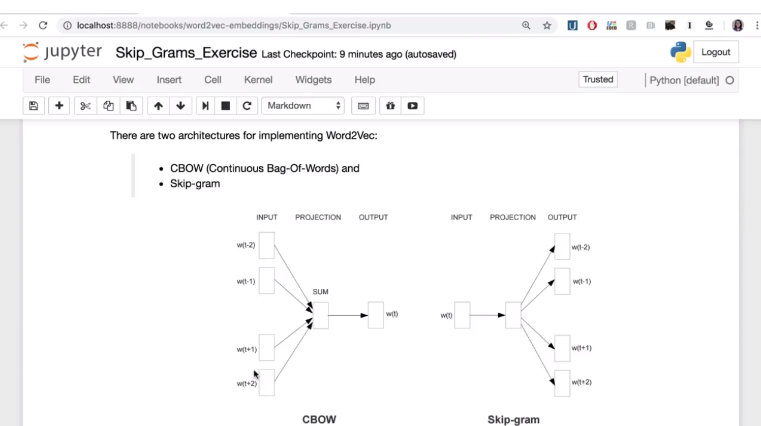


A look up table, 1k words the blue is a vector representation of that word. Embedding dimention

Why embedding? Because we can group the similar words tighter. Lets say for the memory we can use it to give the last 4k W mode as a group and give them similar embedding



Coffee is more suitable to working time

Here comes the skip

Skip-gram input is when we input a word and output 5-6 contex why? Cause it works better

Contex is 2 words before and 2 words after = for memory we can use the last 10 digit as a contex :D

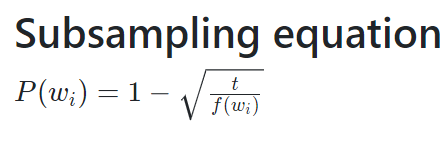
Current time is 18:26:20

Today is Sunday, 7/7/2019

Counter will return words and their frequency: 12412343 , 235123 using these words to build a Dictionary

Most frequent word “The” is set to zero since it has the most frequency then The can be tokenized as 0

Our goal is implemented word2vec context similar like Obama and trump {President}

**Subsampling**

We can omit the words that are once in a while with the above formula we can omit the lower frequency words 60k > 12k vocab

So the address 123456, 303 was used 303 times