

## 输入词向量集合

### One-hot Encoding

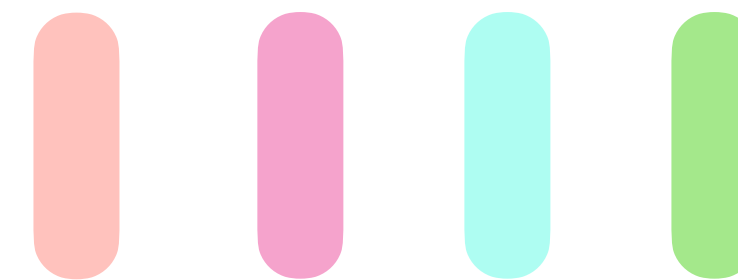
this = [ 1, 0, 0, 0 ]

is = [ 0, 1, 0, 0 ]

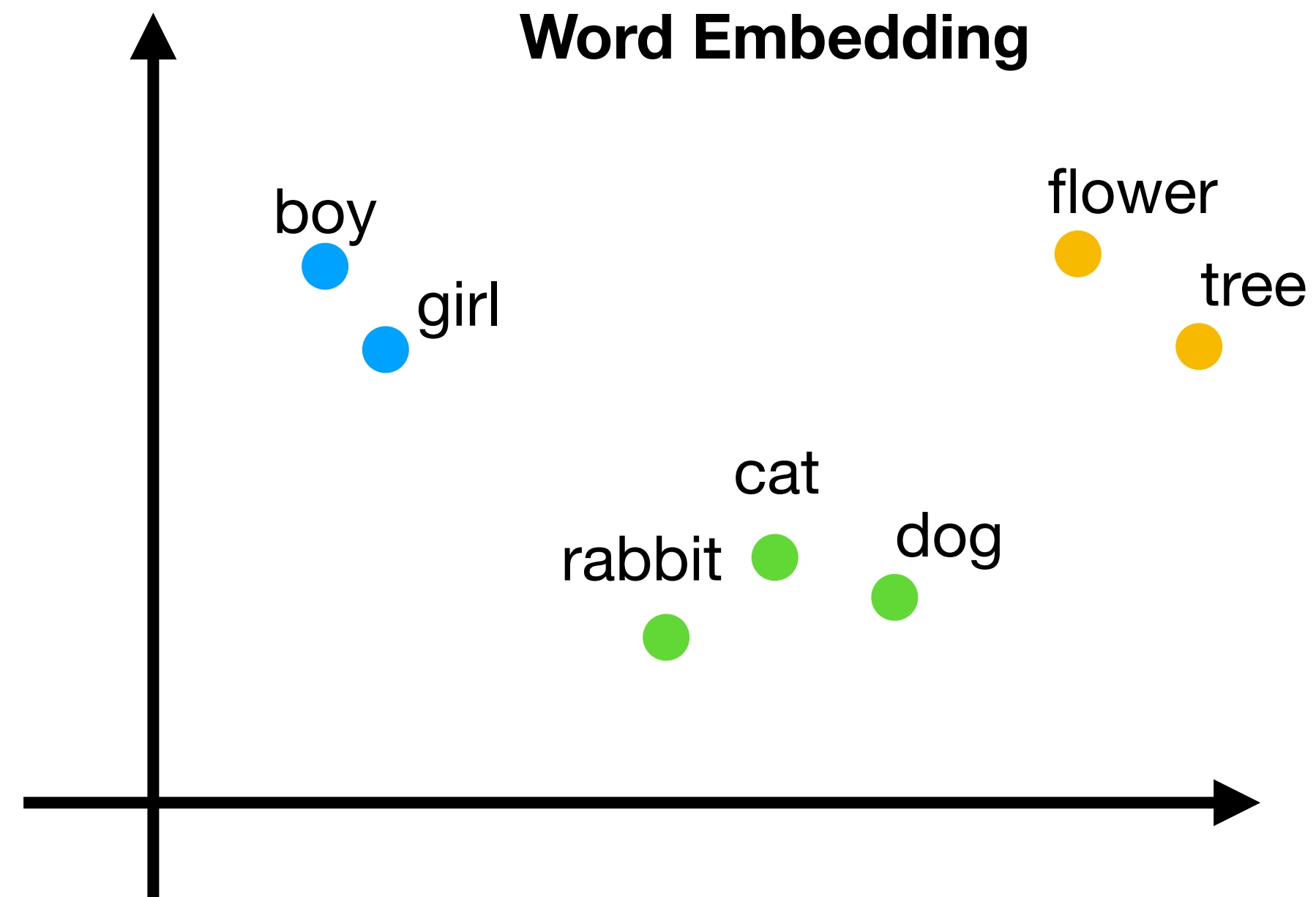
a = [ 0, 0, 1, 0 ]

cat = [ 0, 0, 0, 1 ]

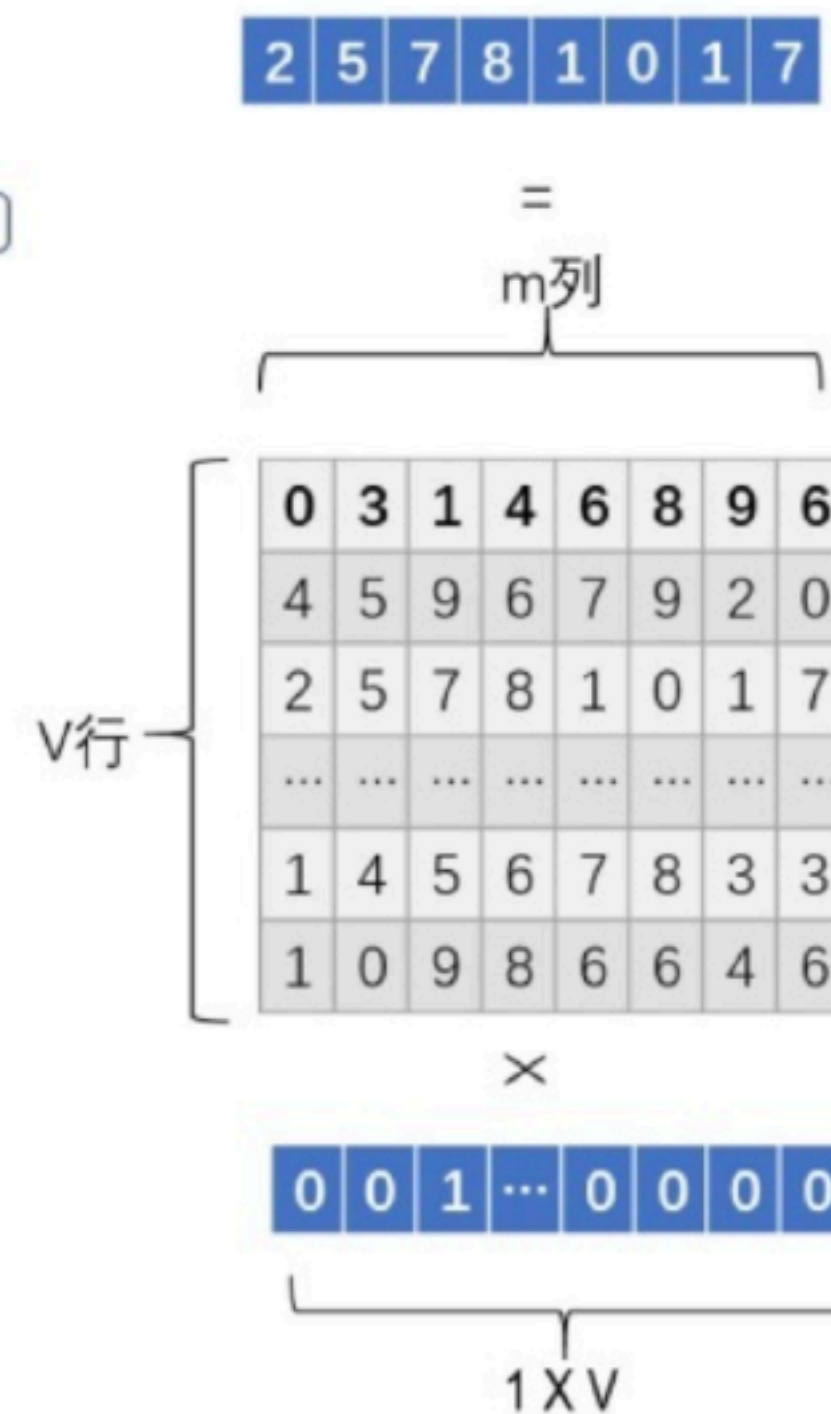
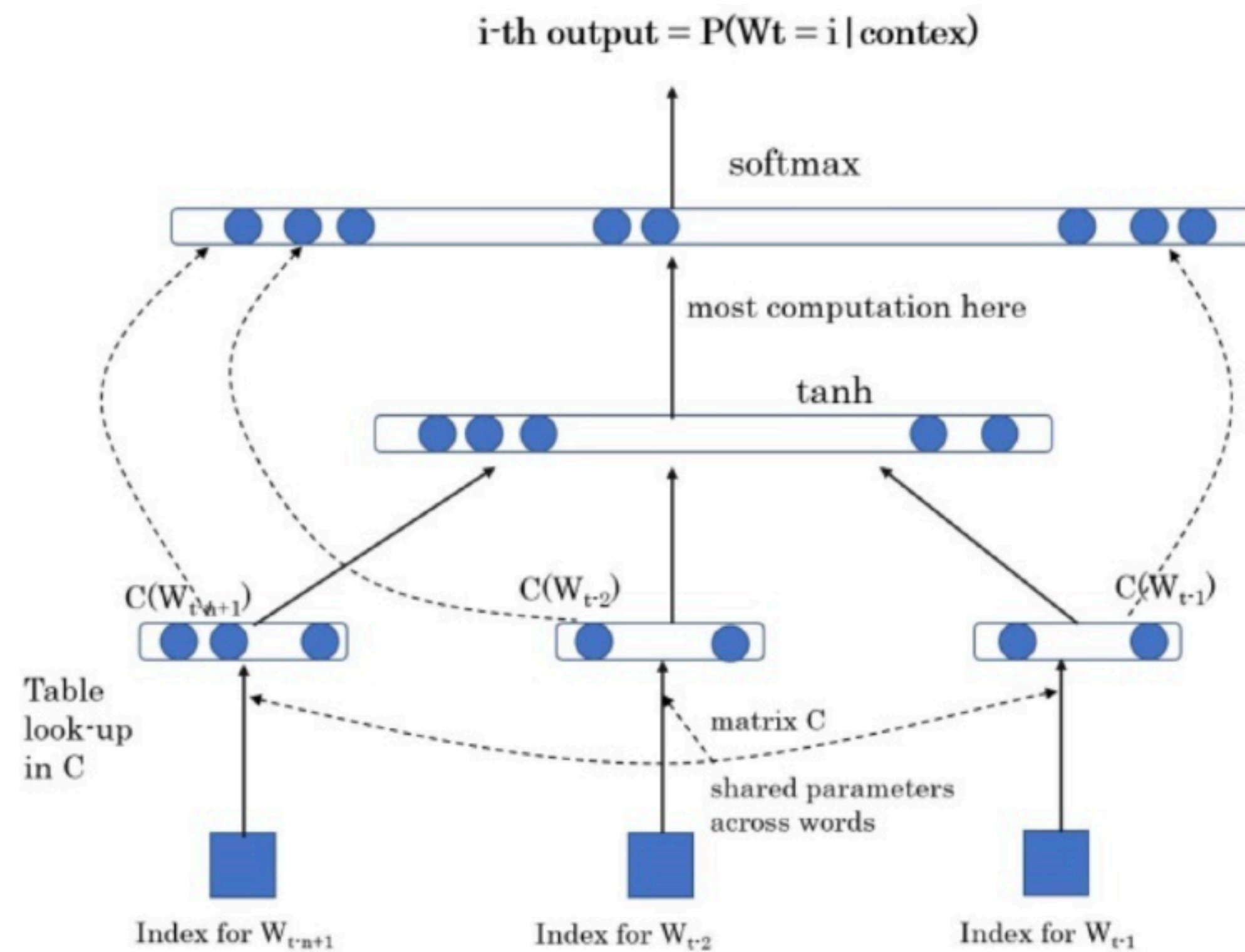
this is a cat



### Word Embedding



# NNLM(Neural Network Language Models)



Input Layer: word indices  
 Projection Layer: word embedding  
 Hidden Layer: vector calculation  
 Output Layer: the probability or the next word

**Word Embedding**

$$y = \text{softmax}(b + Wx + U \tanh(d + Hx))$$