

机器要怎么做？

圆圆爱吃瓜

如何让机器认识一句话中的每个词

圆圆 -> (1, 0, 0, 0)
爱 -> (0, 1, 0, 0)
吃 -> (0, 0, 1, 0)
瓜 -> (0, 0, 0, 1)

Vocabulary_Size = 4

编码特点：

- 1. 词典大小固定
- 2. 损失位置信息

ONE-HOT Encoding

	圆圆	爱	吃	瓜
圆圆	1	0	0	0
爱	0	1	0	0
吃	0	0	1	0
瓜	0	0	0	1

机器要怎么做？



Pic-Embedding

feature1 = (小女孩, 夕阳, 黄色调 ..., 花, 面向)
feature2 = (小女孩, 夕阳, 黄色调 ..., 花, 侧向)
feature2 = (小女孩, 夕阳, 黄色调 ..., 花, 花环)
feature2 = (小女孩, 夕阳, 黄色调 ..., 花, 湖)

one-hot encoding

	面向	侧向	花环	湖
面向	1	0	0	0
侧向	0	1	0	0
花环	0	0	1	0
湖	0	0	0	1

2Vec

vec1 = (... , 1, 0, 0, 0)
vec2 = (... , 0, 1, 0, 0)
vec3 = (... , 0, 0, 1, 0)
vec4 = (... , 0, 0, 0, 1)