

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)

Word Embedding

In natural language processing (NLP), a word embedding is a representation of a word. The embedding is used in text analysis.

Typically, the representation is a real-valued vector that encodes the meaning of the word in such a way that words that are closer in the vector space are expected to be similar in meaning.

用向量表示词汇，将词汇映射到向量空间，在向量空间中，特征相近的词汇靠的会比较近

<https://projector.tensorflow.org/>

1w points, 200 dim

