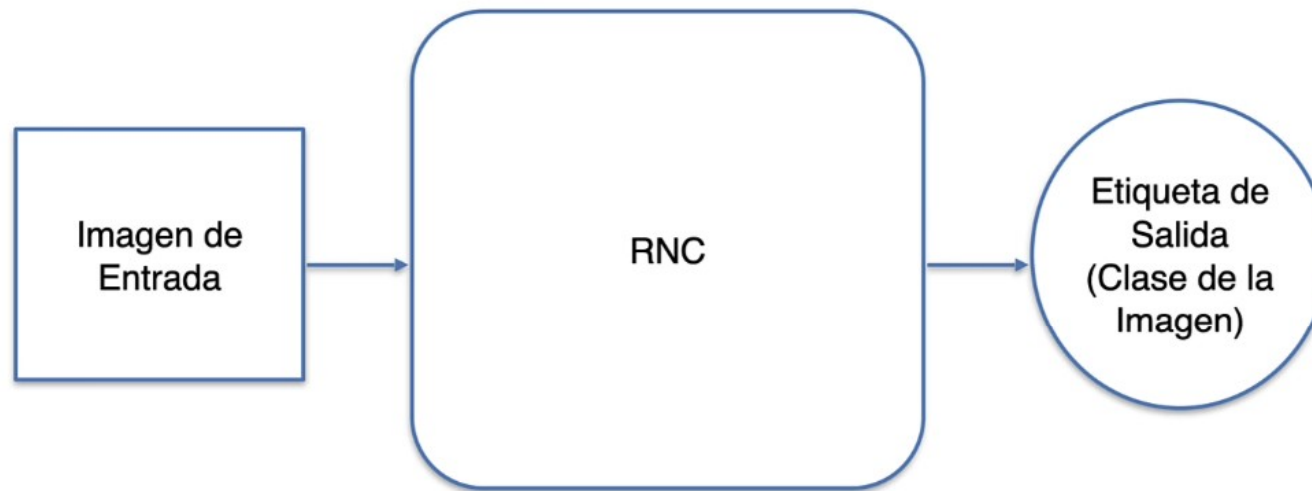


RNC para imágenes



RNC para imágenes

PASO 1: La convolución crea "detectores de características" que recorren la imagen



PASO 2: Max Pooling

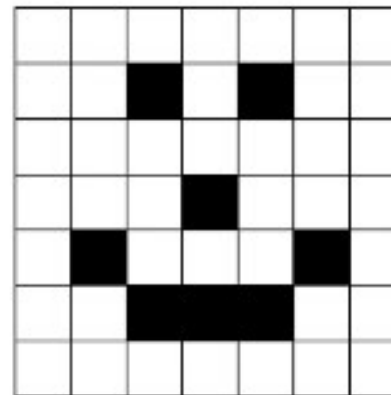


PASO 3: Aplanado



PASO 4: Conexión completa

PASO 1: Convolución



0	0	0	0	0	0	0
0	1	0	0	0	1	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	1	0	0	0	1	0
0	0	1	1	1	0	0
0	0	0	0	0	0	0

PASO 1: Convolución

0	0	0	0	0	0	0
0	1	0	0	0	1	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	1	0	0	0	1	0
0	0	1	1	1	0	0
0	0	0	0	0	0	0

Imagen de
Entrada



0	0	1
1	0	0
0	1	1

Detector de
Características



0				

Mapa de
Características

PASO 1: Convolución

0	0	0	0	0	0	0
0	1	0	0	0	1	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	1	0	0	0	1	0
0	0	1	1	1	0	0
0	0	0	0	0	0	0

Imagen de
Entrada



0	0	1
1	0	0
0	1	1

Detector de
Características



0	1			

Mapa de
Características

PASO 1: Convolución

0	0	0	0	0	0	0
0	1	0	0	0	1	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	1	0	0	0	1	0
0	0	1	1	1	0	0
0	0	0	0	0	0	0

Imagen de
Entrada



0	0	1
1	0	0
0	1	1

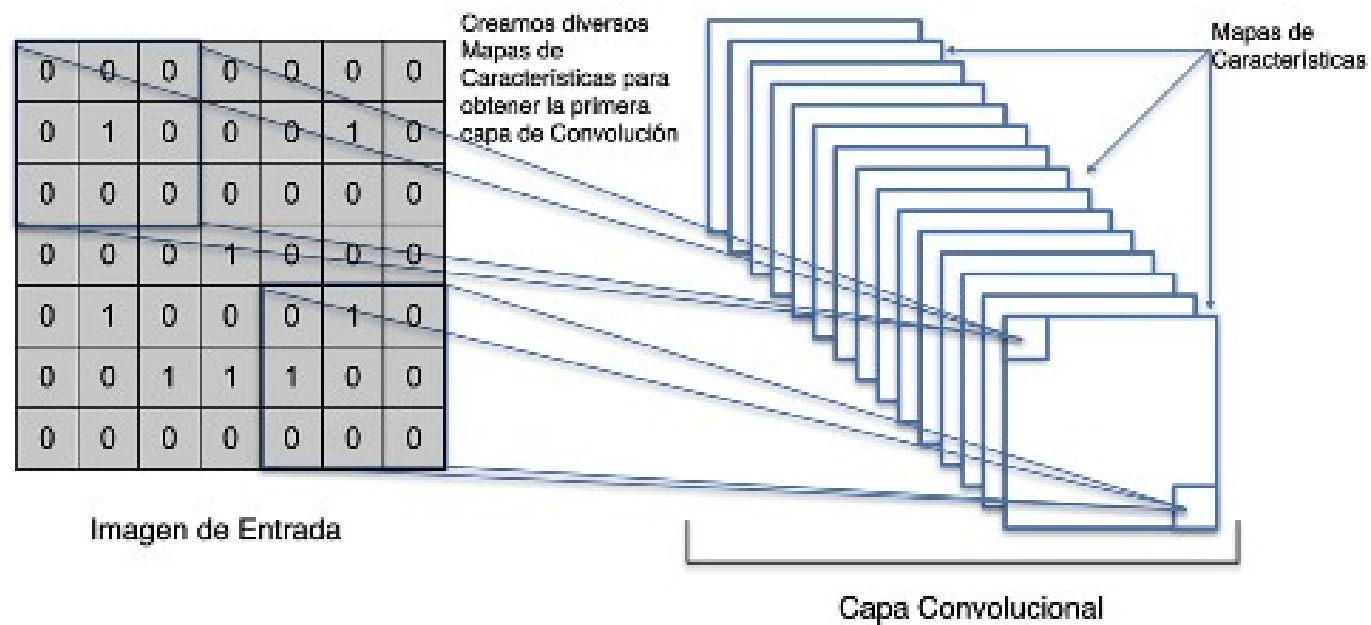
Detector de
Características



0	1	0	0	0
0	1	1	1	0
1	0	0	2	1
1	4	2	1	0
0	0	1	2	

Mapa de
Características

PASO 1: Convolución



RNC para imágenes

PASO 1: La convolución crea "detectores de características" que recorren la imagen



PASO 2: Max Pooling: reduce el tamaño y el coste, convirtiéndolo en más genérico



PASO 3: Aplanado



PASO 4: Conexión completa

PASO 2: max pooling

0	1	0	0	0
0	1	1	1	0
1	0	1	2	1
1	4	2	1	0
0	0	1	2	1

Mapa de
Características

Max Pooling

1		

Mapa de
Características
Pooled

PASO 2: max pooling

0	1	0	0	0
0	1	1	1	0
1	0	1	2	1
1	4	2	1	0
0	0	1	2	1

Mapa de
Características

Max Pooling

1	1	

Mapa de
Características
Pooled

PASO 2: max pooling

0	1	0	0	0
0	1	1	1	0
1	0	1	2	1
1	4	2	1	0
0	0	1	2	1

Mapa de
Características

Max Pooling

1	1	0

Mapa de
Características
Pooled

PASO 2: max pooling

0	1	0	0	0
0	1	1	1	0
1	0	1	2	1
1	4	2	1	0
0	0	1	2	1

Mapa de
Características

Max Pooling

1	1	0
4		

Mapa de
Características
Pooled

PASO 2: max pooling

0	1	0	0	0
0	1	1	1	0
1	0	1	2	1
1	4	2	1	0
0	0	1	2	1

Mapa de
Características

Max Pooling

1	1	0
4	2	1
0		

Mapa de
Características
Pooled

PASO 2: max pooling

0	1	0	0	0
0	1	1	1	0
1	0	1	2	1
1	4	2	1	0
0	0	1	2	1

Mapa de
Características

Max Pooling

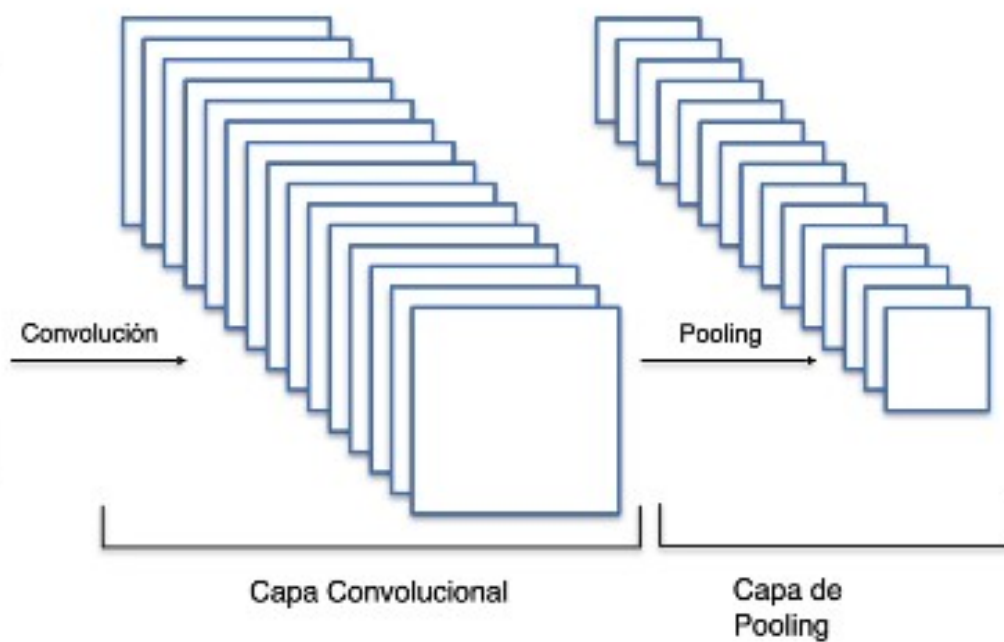
1	1	0
4	2	1
0	2	1

Mapa de
Características
Pooled

RNC para imágenes

0	0	0	0	0	0	0
0	1	0	0	0	1	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	1	0	0	0	1	0
0	0	1	1	1	0	0
0	0	0	0	0	0	0

Imagen de
Entrada



RNC para imágenes

PASO 1: La convolución crea "detectores de características" que recorren la imagen



PASO 2: Max Pooling: reduce el tamaño y el coste, convirtiéndolo en más genérico



PASO 3: Aplanado: convertir todo a un solo vector unidimensional



PASO 4: Conexión completa

RNC para imágenes

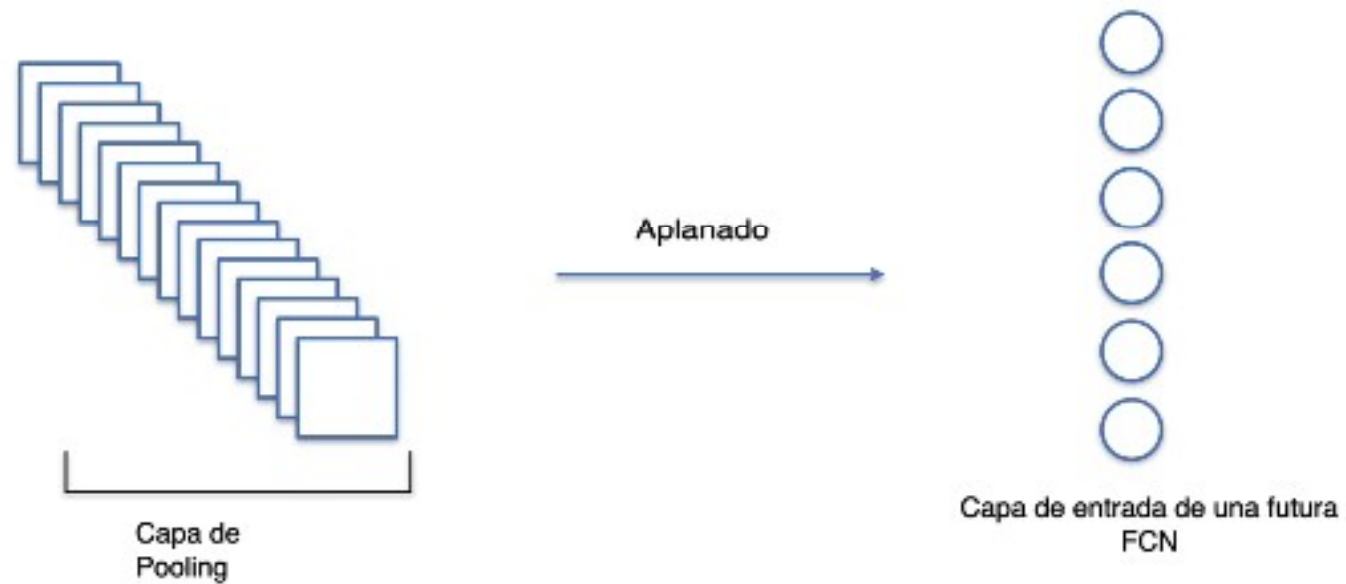
1	1	0
4	2	1
0	2	1

Mapa de
Características
Pooled

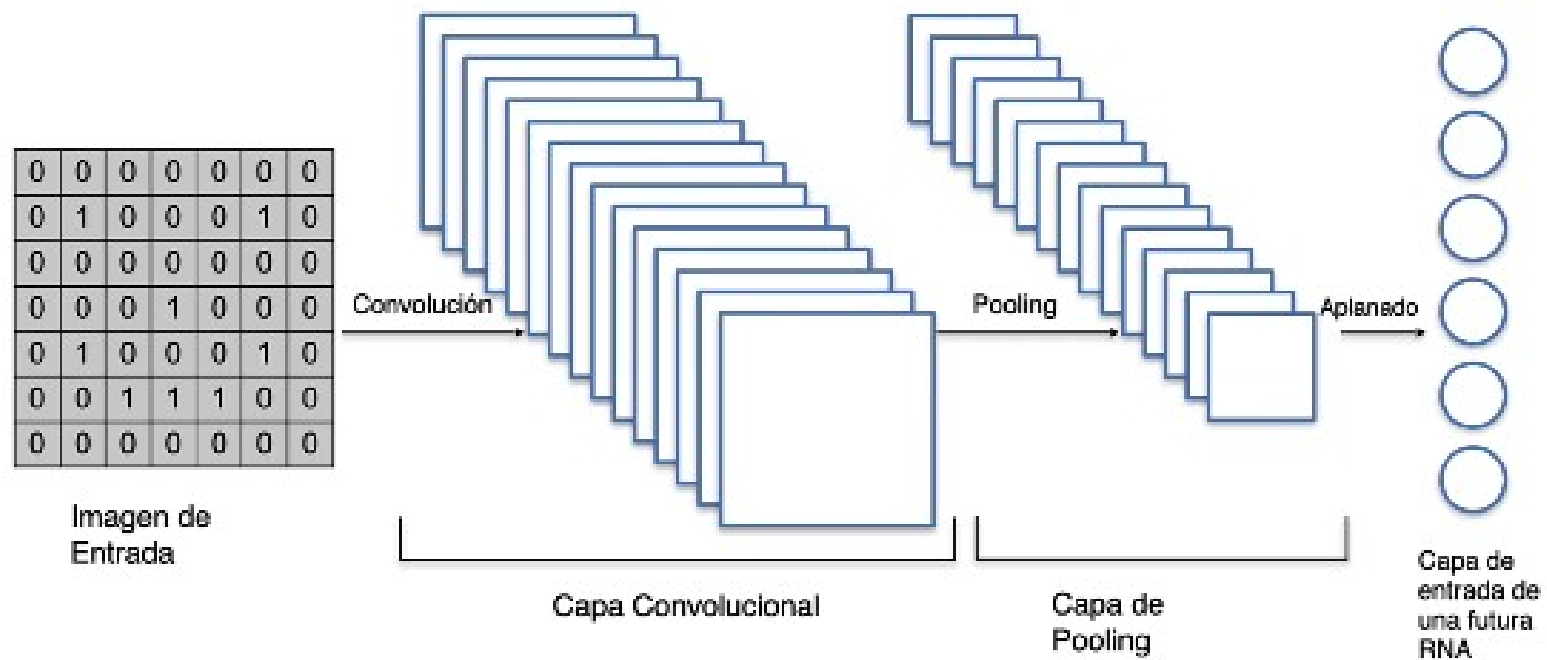
Aplanado

1
1
0
4
2
1
0
2
1

PASO 3: Aplanado



RNC para imágenes



RNC para imágenes

PASO 1: La convolución crea "detectores de características" que recorren la imagen



PASO 2: Max Pooling: reduce el tamaño y el coste, convirtiéndolo en más genérico

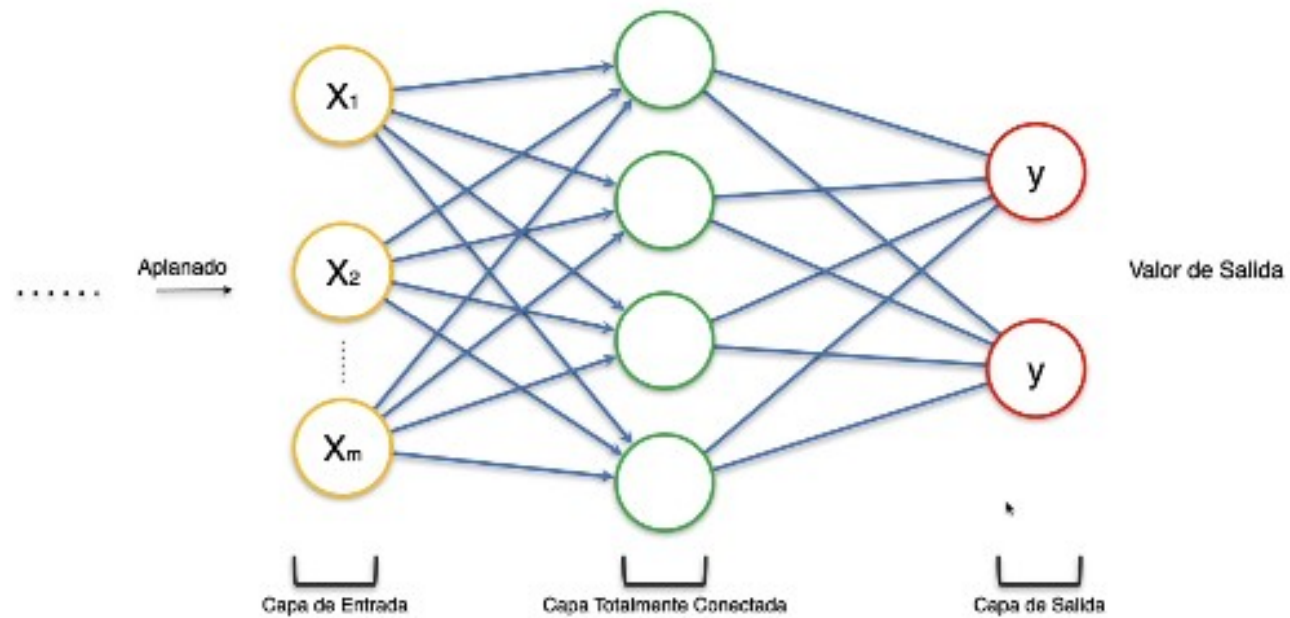


PASO 3: Aplanado: convertir todo a un solo vector unidimensional

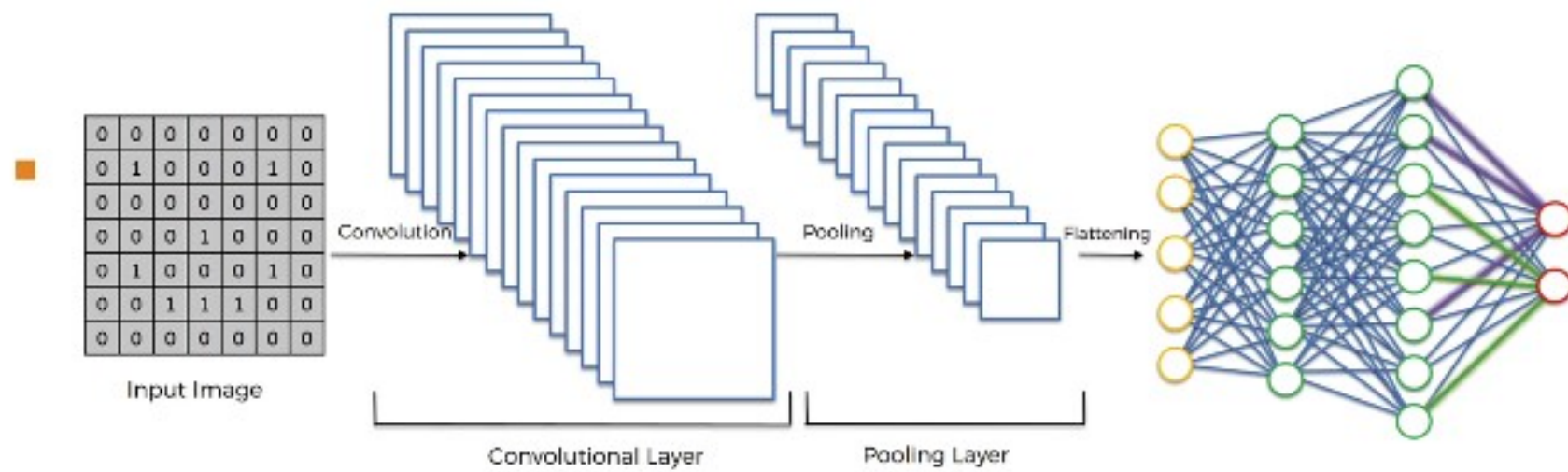


PASO 4: Conexión completa: RNA totalmente conectada para aprender a clasificar

RNC para imágenes



RNC para imágenes



RNC - Teoría

De la imagen al texto





De imagen a texto

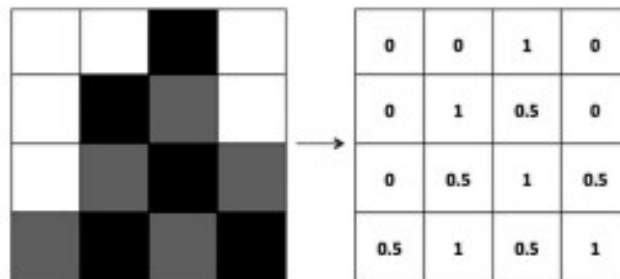


RNC para imágenes: buscamos características locales en una imagen. ¿Por qué no probar lo mismo para frases de texto?



De imagen a texto

Representación de las imágenes



Representación del texto



Word embedding

Representación sencilla pero inefectiva de palabras: one-hot encoding. No hay relación entre las palabras.

$$\text{dog} = [0, \dots, 0, \overset{\text{vocab_size}}{\underbrace{1}_{\substack{\uparrow \\ \text{dog}}}}, 0, \dots, 0] \quad \text{vocab_size} \approx 100,000$$

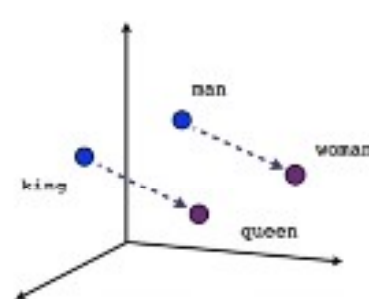
Word embedding

Word embedding: hacemos cada vector más pequeño \Rightarrow añadimos relaciones entre palabras.

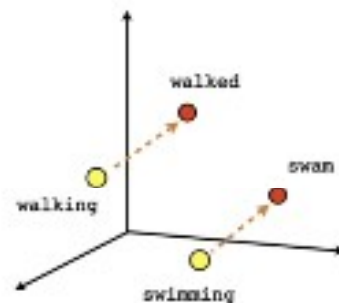
$$\text{dog} = \overbrace{[0.194, 0.047, \dots, 0.126]}^{emb_dim} \quad emb_dim \approx 64$$

Word embedding

Word embedding: relación matemática entre las palabras



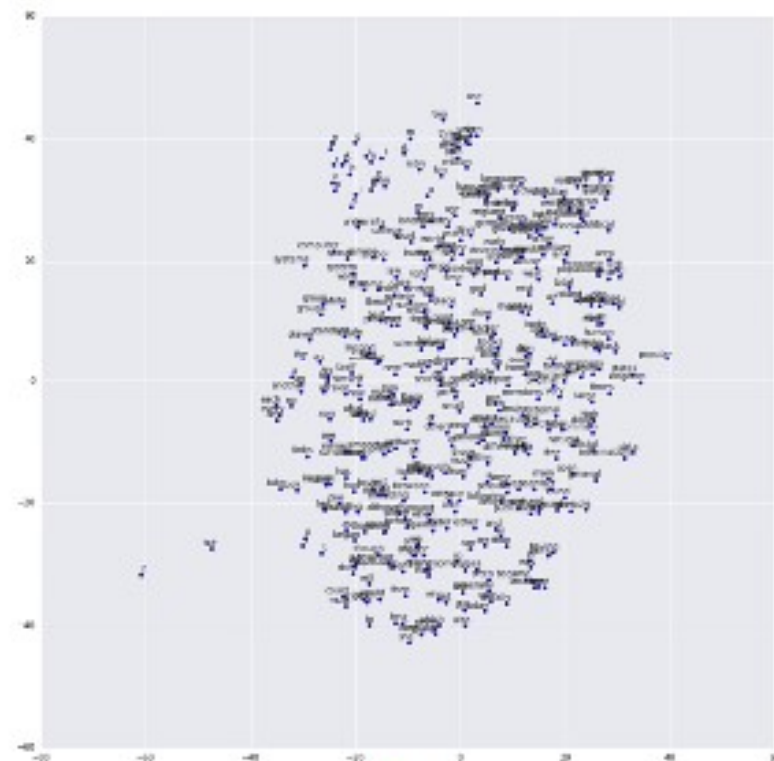
Male-Female



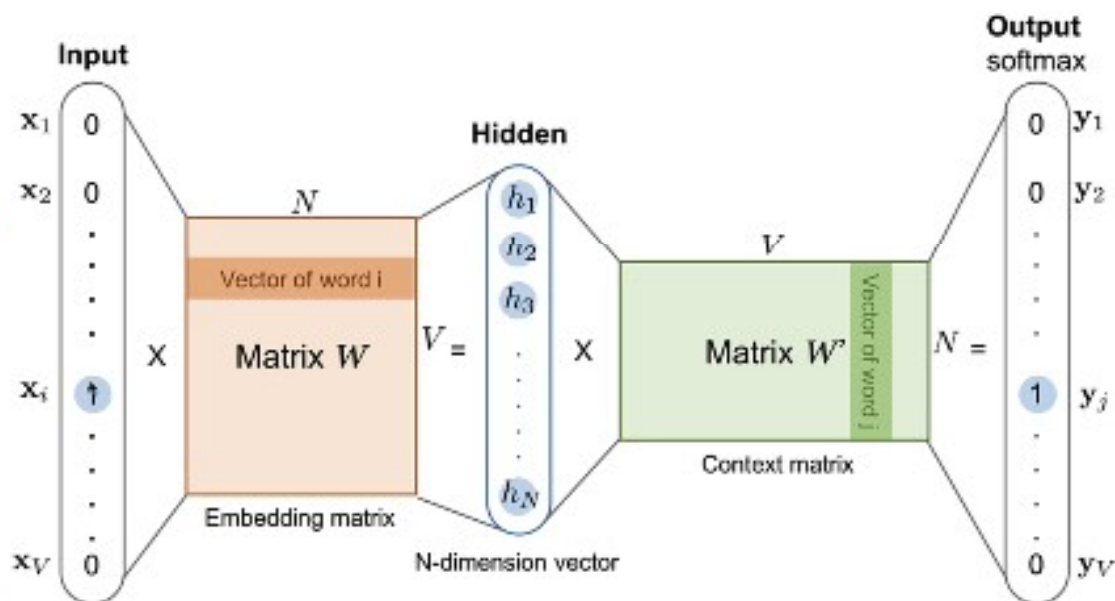
Verb tense

$$[\text{king}] - [\text{man}] + [\text{woman}] = [\text{queen}]$$

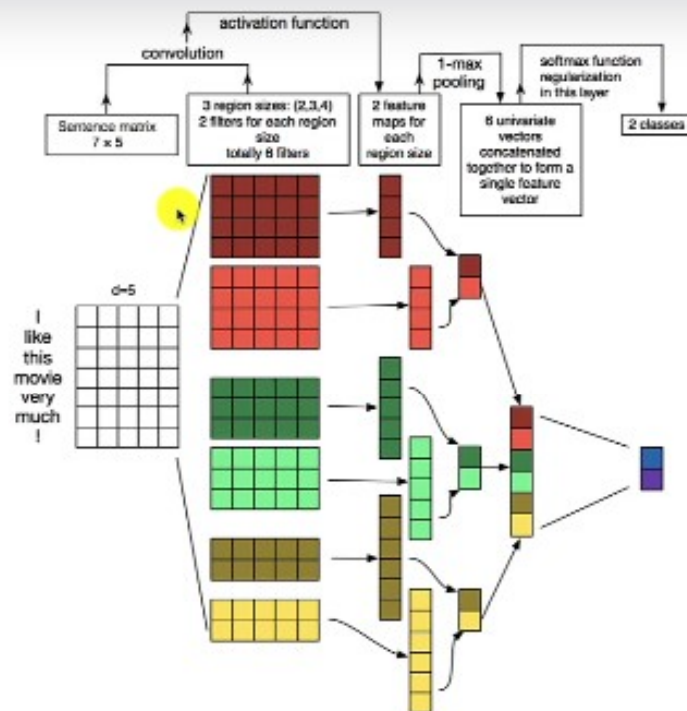
$$[\text{Paris}] - [\text{France}] + [\text{Italia}] = [\text{Rome}]$$



Word embedding



Modelo Skip-grama: en la frase "In spite of everything, I still believe people are really good at heart", la palabra "good" produce los pares ("good", "are"), ("good", "really"), ("good", "at"), ("good", "heart") en forma de target/contexto.



Diferencias Principales:

- Cada filtro tiene anchura = d_{model} . Dividir la dimensión de embedding no tiene sentido.
- Elegimos el máximo para cada filtro. La posición de la característica en la frase es menos relevante.
- 3 tamaños diferentes de filtros, para capturar diferentes niveles de correlación entre las palabras.