







Cummins

Text  
Classification  
Model

Person

Country



How can we capture  
similarities between  
two words?



Bedrooms: 3  
Area: 1850 sqr ft  
Bathrooms: 2.5



Bedrooms: 3  
Area: 1700 sqr ft  
Bathrooms: 2



Bedrooms: 10  
Area: 7500 sqr ft  
Bathrooms: 2



Dhoni

Person: 1  
 Healthy/fit: 0.9  
 Location: 0  
 Has two eyes: 1  
 Has government: 0



Cummins

Person: 1  
 Healthy/fit: 0.87  
 Location: 0  
 Has two eyes: 1  
 Has government: 0



Australia

Person: 0  
 Healthy/fit: 0.7  
 Location: 1  
 Has two eyes: 0  
 Has government: 1

Image credits: cnn.com, Britannica.com and outlookindia.com





Dhoni

$$\begin{bmatrix} 1 \\ 0.9 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$



Cummins

$$\begin{bmatrix} 1 \\ 0.87 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$



Australia

$$\begin{bmatrix} 0 \\ 0.7 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Image credits: cnn.com, Britannica.com and outlookindia.com

	ashes	Australia	Bat	Cummins	cover	Dhoni	World cup	..	Zimbabwe
Person	0	0.02	0.1	0.95	0.03	0.96	0.67	...	0.04
Country	0	0.97	0	0	0	0	0	...	1
Healthy & Fit	0	0	0.3	0.87	0	0.9	0	...	0
Event	1	0.1	0	0	0.4	0	1	...	0
gear	0	0	1	0	0	0	0	...	0



	ashes	Australia	Bat	Cummins	cover	Dhoni	World cup	..	Zimbabwe
Person	0	0.02	0.1	0.95	0.03	0.96	0.67	...	0.04
Country	0	0.97	0	0	0	0	0	...	1
Healthy & Fit	0	0	0.3	0.87	0	0.9	0	...	0
Event	1	0.1	0	0	0.4	0	1	...	0
gear	0	0	1	0	0	0	0	...	0

Cummins	0.95	0.96	Dhoni	Australia	0.02	0.04	Zimbabwe
	0	0			0.97	1	
	1	1			0	0	
	0	0			0.1	0	
	0	0			0	0	

# Issue with word2vec

## Fixed embeddings

He didn't receive fair treatment



Fun fair in new York city this summer

# Issue with word2vec

## Fixed embeddings

He didn't receive fair treatment

$$\text{fair} \rightarrow \begin{bmatrix} 1 \\ 0.9 \\ 0.2 \\ 1 \\ 0.7 \end{bmatrix}$$

Fun fair in new York city this summer

# BERT can generate contextualized embeddings



# BERT can generate contextualized embeddings



# BERT can generate embeddings for entire sentence

Amazing movie, I couldn't blink an eye for initial 45 minutes. It was that intense and interesting at the same time

$$\begin{bmatrix} 1 \\ 0.9 \\ 0.2 \\ 1 \\ 0.7 \\ \dots \\ 0.1 \end{bmatrix}$$