



Dhoni



Cummins



Australia

Person: 1

Healthy/fit: 0.9

Location: 0

Has two eyes: 1

Has government: 0

Person: 1

Healthy/fit: 0.87

Location: 0

↳ Has two eyes: 1

Has government: 0

Person: 0

Healthy/fit: 0.7

Location: 1

Has two eyes: 0

Has government: 1

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



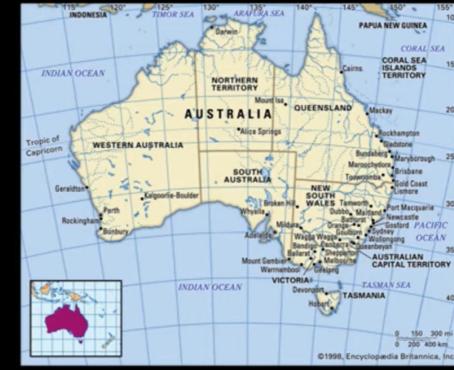
Dhoni

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



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



Embeddings are not hand
crafted. Instead, they are
learnt during neural network
training

Techniques to compute word embeddings

1. Using **supervised** learning
2. Using **self-supervised** learning
 1. Word2vec
 2. Glove

1. Using **supervised** learning

Take an NLP problem and try to solve it.
In that pursuit as a side effect, you get
word embeddings

nice food. The pasta dish was too good!



poor quality food. I would never go there again!



After food nice ... poor ... zonal
1 2 3 5000

$$\begin{bmatrix} 0.5 \\ 1.2 \\ 0.7 \\ 3.1 \end{bmatrix} \quad \begin{bmatrix} 4.3 \\ 0.1 \\ 0.9 \\ 5.5 \end{bmatrix} \quad \begin{bmatrix} 0.4 \\ 8.1 \\ 8.8 \\ 4.2 \end{bmatrix}$$

$$\begin{bmatrix} 9.8 \\ 0.6 \\ 2.2 \\ 1.3 \end{bmatrix}$$



4-dimensional word embedding vector

After	food	nice	...	poor	...	zonal
1	2	3				5000
0.5	4.3	0.4				9.8
1.2	0.1	8.1				0.6
0.7	0.9	8.8	...			2.2
3.1	5.5	4.2				1.3

4 by 5000 Embedding Matrix: E

	After	food	nice	...	poor	...	zonal
After	1	0	0		0		0
food	0	1	0		0		0
nice	0	0	1	*	0		0
...							
poor	0	0	0		1		0
...							
zonal	0	0	0		0		1

Training sample 1

After	food	nice	...	poor	...	zonal
1	2	3				5000

nice

0.5	0.8	0.9				0.1
0.2	0.6	0.2				0.8
0.3	0.3	0.1	...			0.1
0.4	0.1	0.7				0.8

After	food	nice	...	poor	...	zonal
1	2	3				5000

food

0.5	0.8	0.9				0.1
0.2	0.6	0.2				0.8
0.3	0.3	0.1	...			0.1
0.4	0.1	0.7				0.8

Training sample 1

After food nice ... poor ... zonal

1

2
3

0.5	0.8	0.9
0.2	0.6	0.2
0.3	0.3	0.1
0.4	0.1	0.7

...

poor ...

zonal

5000

0.1

0.8

0.1

0.8

X

O	O	1	O	...	O
---	---	---	---	-----	---

nice



After food nice ... poor ... zonal

1

2
3

0.5	0.8	0.9
0.2	0.6	0.2
0.3	0.3	0.1
0.4	0.1	0.7

...

poor ...

zonal

5000

0.1

0.8

0.1

0.8

X

O	1	O	O	...	O
---	---	---	---	-----	---

food

Training sample 1

nice

After	food	nice	...	poor	...	zonal
1	2	3				5000
0.5	0.8	0.9				0.1
0.2	0.6	0.2				0.8
0.3	0.3	0.1				0.1
0.4	0.1	0.7	0.8

$$\begin{bmatrix} 0 \\ 0 \\ \mathbf{1} \\ 0 \\ \dots \\ 0 \end{bmatrix} \times \rightarrow \begin{bmatrix} 0.9 \\ 0.2 \\ 0.1 \\ 0.7 \end{bmatrix}$$

food

After	food	nice	...	poor	...	zonal
1	2	3				5000
0.5	0.8	0.9				0.1
0.2	0.6	0.2				0.8
0.3	0.3	0.1				0.1
0.4	0.1	0.7	0.8

$$\begin{bmatrix} 0 \\ \mathbf{1} \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \times \rightarrow \begin{bmatrix} 0.8 \\ 0.6 \\ 0.3 \\ 0.1 \end{bmatrix}$$

Training sample 1

nice

After	food	nice	...	poor	...	zonal
1	2	3				5000
0.5	0.8	0.9				0.1
0.2	0.6	0.2				0.8
0.3	0.3	0.1				0.1
0.4	0.1	0.7	0.8

$$\begin{bmatrix} O \\ O \\ \mathbf{1} \\ O \\ \dots \\ O \end{bmatrix} \times \rightarrow \begin{bmatrix} 0.9 \\ 0.2 \\ 0.1 \\ 0.7 \end{bmatrix}$$

$$\begin{bmatrix} 0.9 \\ 0.2 \\ 0.1 \\ 0.7 \end{bmatrix} \rightarrow \begin{bmatrix} 0.9 \\ 0.2 \\ 0.1 \\ 0.7 \end{bmatrix}$$

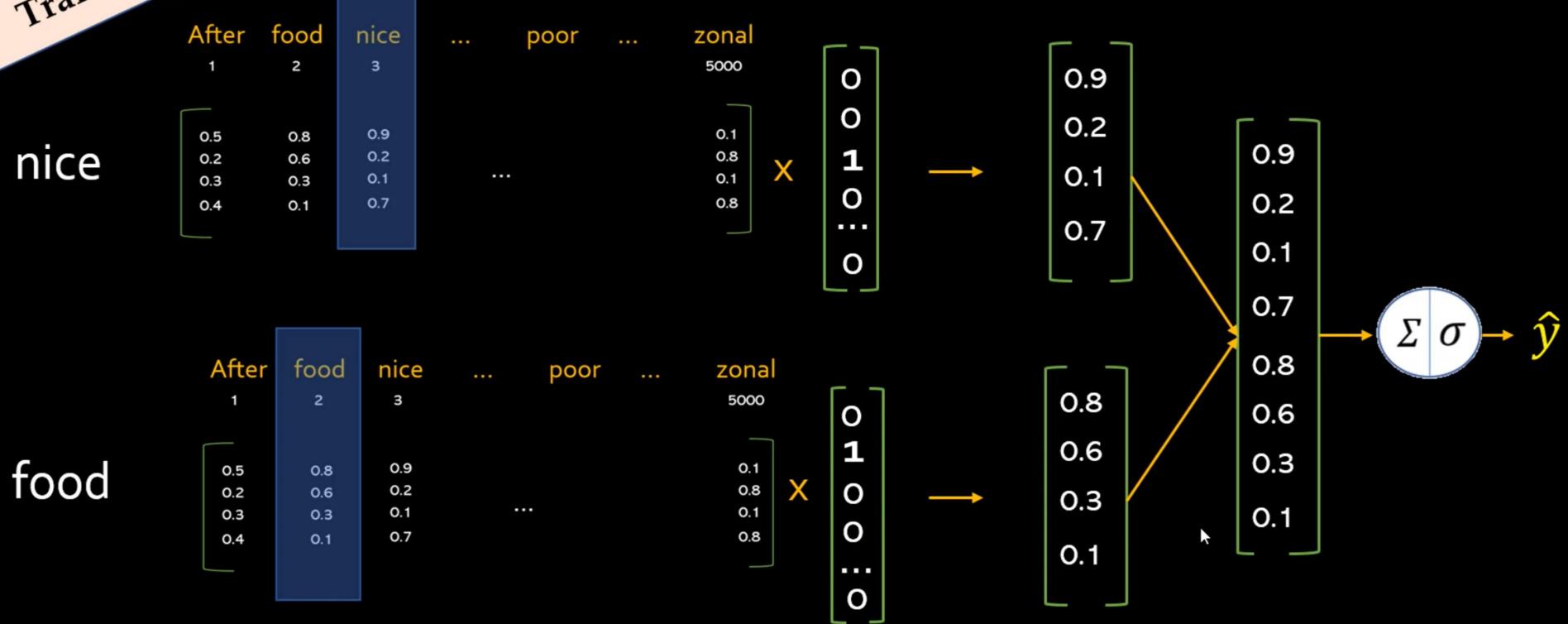
food

After	food	nice	...	poor	...	zonal
1	2	3				5000
0.5	0.8	0.9				0.1
0.2	0.6	0.2				0.8
0.3	0.3	0.1				0.1
0.4	0.1	0.7	0.8

$$\begin{bmatrix} O \\ \mathbf{1} \\ O \\ O \\ \dots \\ O \end{bmatrix} \times \rightarrow \begin{bmatrix} 0.8 \\ 0.6 \\ 0.3 \\ 0.1 \end{bmatrix}$$

$$\begin{bmatrix} 0.8 \\ 0.6 \\ 0.3 \\ 0.1 \end{bmatrix} \rightarrow \begin{bmatrix} 0.8 \\ 0.6 \\ 0.3 \\ 0.1 \end{bmatrix}$$

Training sample 1



Training sample 1

nice

After	food	nice
1	2	3
0.5	0.8	0.9
0.2	0.6	0.2
0.3	0.3	0.1
0.4	0.1	0.7

...

poor

...

zonal

5000
0.1
0.8
0.1
0.8

$$\times \begin{bmatrix} O \\ O \\ 1 \\ O \\ O \end{bmatrix}$$



Back propagation

food

After	food	nice
1	2	3
0.5	0.8	0.9
0.2	0.6	0.2
0.3	0.3	0.1
0.4	0.1	0.7

...

poor

...

zonal

5000
0.1
0.8
0.1
0.8

$$\times \begin{bmatrix} O \\ 1 \\ O \\ O \\ O \end{bmatrix}$$

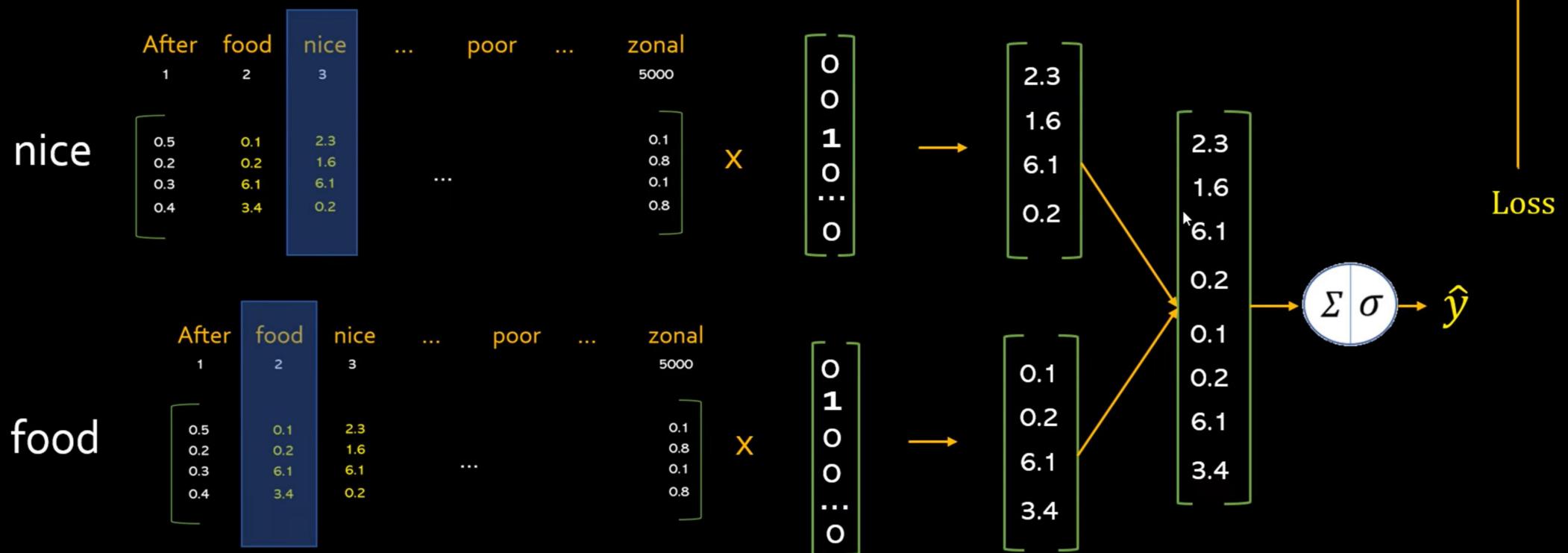


$$\begin{bmatrix} 0.9 \\ 0.2 \\ 0.1 \\ 0.7 \end{bmatrix} \quad \begin{bmatrix} 0.9 \\ 0.2 \\ 0.1 \\ 0.7 \end{bmatrix}$$

$$\Sigma \sigma \rightarrow \hat{y}$$

Loss

Back propagation



Training sample 2

poor

After	food	nice	...	poor	...	zonal
1	2	3				5000
0.5	0.1	2.3				0.1
0.2	0.2	1.6				0.8
0.3	6.1	6.1				0.1
0.4	3.4	0.2				0.8

X

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

$$\begin{bmatrix} 1.1 \\ 1.4 \\ 4.1 \\ 7.2 \\ \dots \\ 1.1 \end{bmatrix}$$

quality

After	food	nice	...	poor	...	zonal
1	2	3				5000
0.5	0.1	2.3				0.1
0.2	0.2	1.6				0.8
0.3	6.1	6.1				0.1
0.4	3.4	0.2				0.8

X

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

$$\begin{bmatrix} 0.1 \\ 0.2 \\ 0.1 \\ 0.1 \\ 0.2 \\ 0.1 \end{bmatrix}$$

food

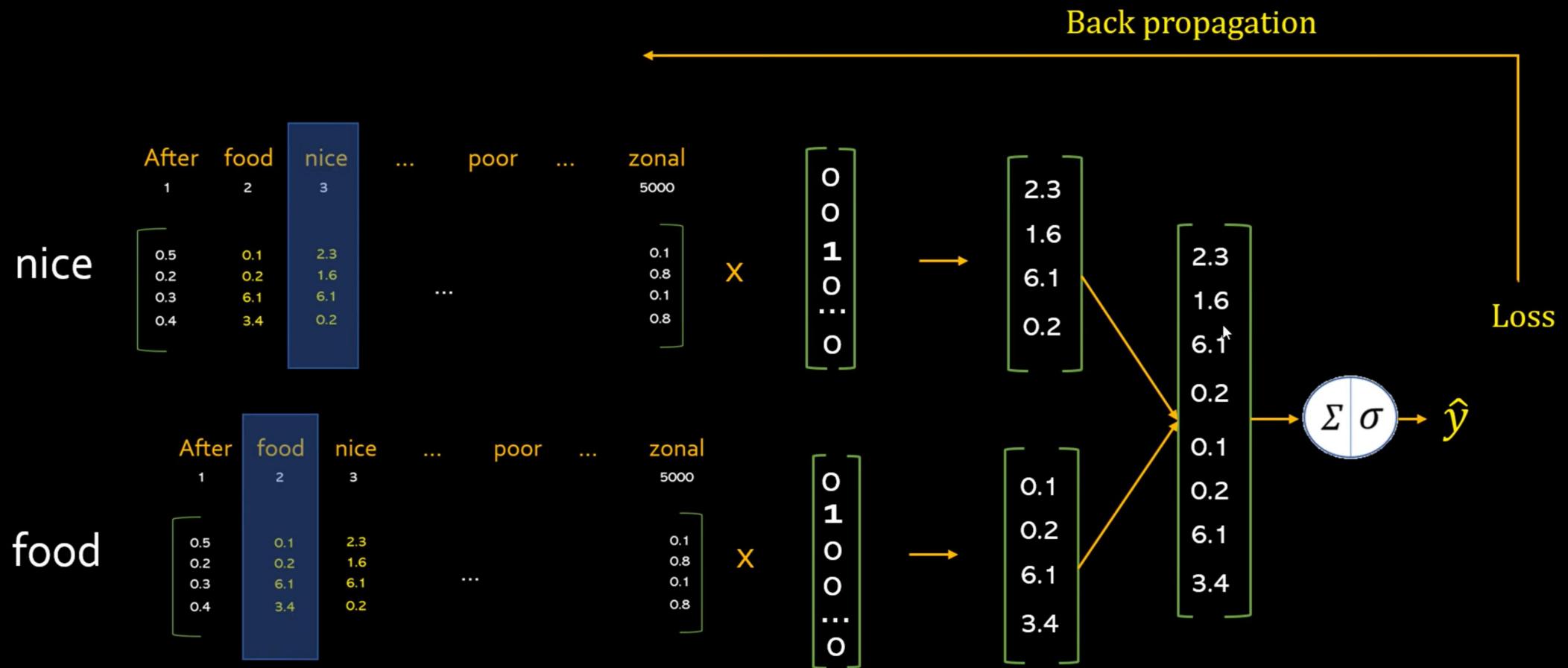
After	food	nice	...	poor	...	zonal
1	2	3				5000
0.5	0.1	2.3				0.1
0.2	0.2	1.6				0.8
0.3	6.1	6.1				0.1
0.4	3.4	0.2				0.8

X

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

$$\begin{bmatrix} 0.1 \\ 0.2 \\ 0.1 \\ 0.1 \\ 0.2 \\ 0.1 \end{bmatrix}$$

$$\Sigma | \sigma \rightarrow \hat{y}$$



Training sample 2

poor

After	food	nice	...	poor	...	zonal
1	2	3				5000
0.5	0.1	2.3				0.1
0.2	0.2	1.6				0.8
0.3	6.1	6.1				0.1
0.4	3.4	0.2				0.8

X

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

$$\begin{bmatrix} 1.1 \\ 1.4 \\ 4.1 \\ 7.2 \\ \dots \\ 1.1 \end{bmatrix}$$

quality

After	food	nice	...	poor	...	zonal
1	2	3				5000
0.5	0.1	2.3				0.1
0.2	0.2	1.6				0.8
0.3	6.1	6.1				0.1
0.4	3.4	0.2				0.8

X

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

$$\begin{bmatrix} 0.1 \\ 0.2 \\ 0.1 \\ 0.1 \\ 0.2 \\ 0.1 \end{bmatrix}$$

food

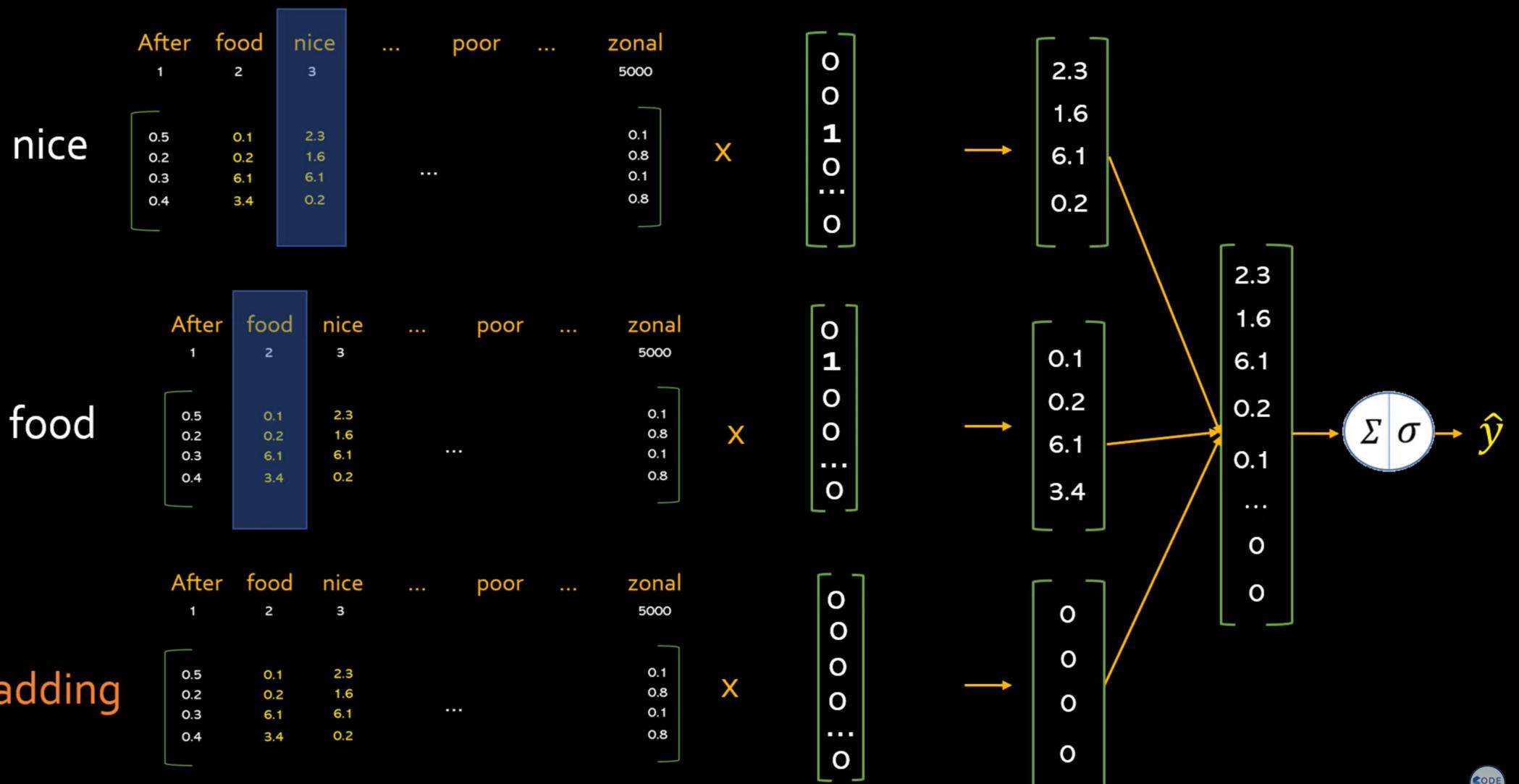
After	food	nice	...	poor	...	zonal
1	2	3				5000
0.5	0.1	2.3				0.1
0.2	0.2	1.6				0.8
0.3	6.1	6.1				0.1
0.4	3.4	0.2				0.8

X

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

$$\begin{bmatrix} 0.1 \\ 0.2 \\ 0.1 \\ 0.1 \\ 0.2 \\ 0.1 \end{bmatrix}$$

$$\Sigma | \sigma \rightarrow \hat{y}$$



After food nice good poor weak ... zonal

5000

0.5	4.3	0.4	0.38	2.3	2.2	9.8
1.2	0.1	8.1	8.2	1.1	1.0	0.6
0.7	0.9	8.8	8.9	6.7	6.5	2.2
3.1	5.5	4.2	4.1	2.0	1.9	1.3

4 by 5000 Embedding Matrix: E

After food nice good poor weak ... zonal

5000

0.5	4.3	0.4	0.38	2.3	2.2	9.8
1.2	0.1	8.1	8.2	1.1	1.0	0.6
0.7	0.9	8.8	8.9	6.7	6.5	2.2
3.1	5.5	4.2	4.1	2.0	1.9	1.3

4 by 5000 Embedding Matrix: E

After food nice good poor weak ... zonal

5000

0.5	4.3	0.4	0.38	2.3	2.2	9.8
1.2	0.1	8.1	8.2	1.1	1.0	0.6
0.7	0.9	8.8	8.9	6.7	6.5	2.2
3.1	5.5	4.2	4.1	2.0	1.9	1.3

4 by 5000 Embedding Matrix: E