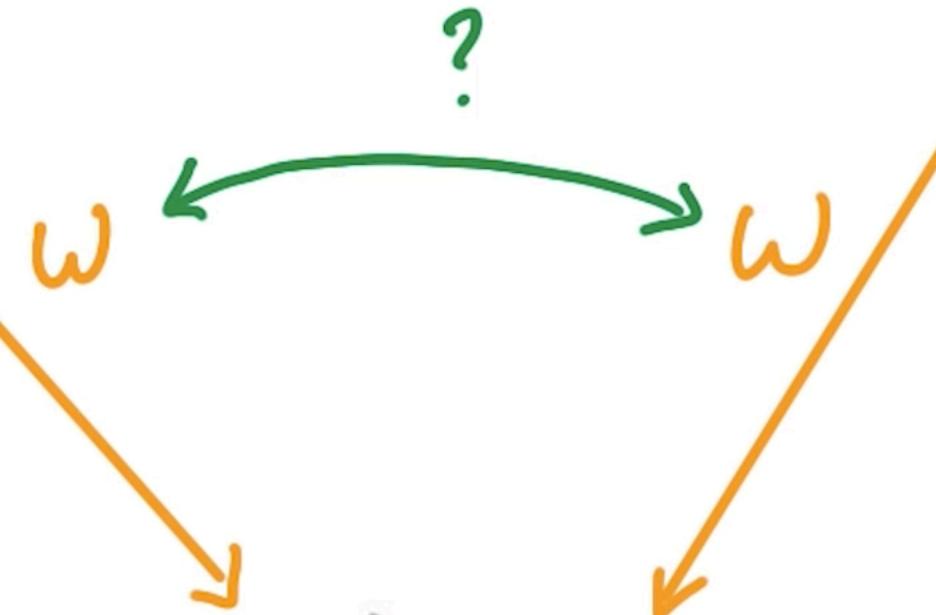


'CAT'

'KITTY'



1)

ret·i·nop·a·thy

/rētn'äpəTHē/

noun MEDICINE

disease of the retina that results in impairment or loss of vision.



2)

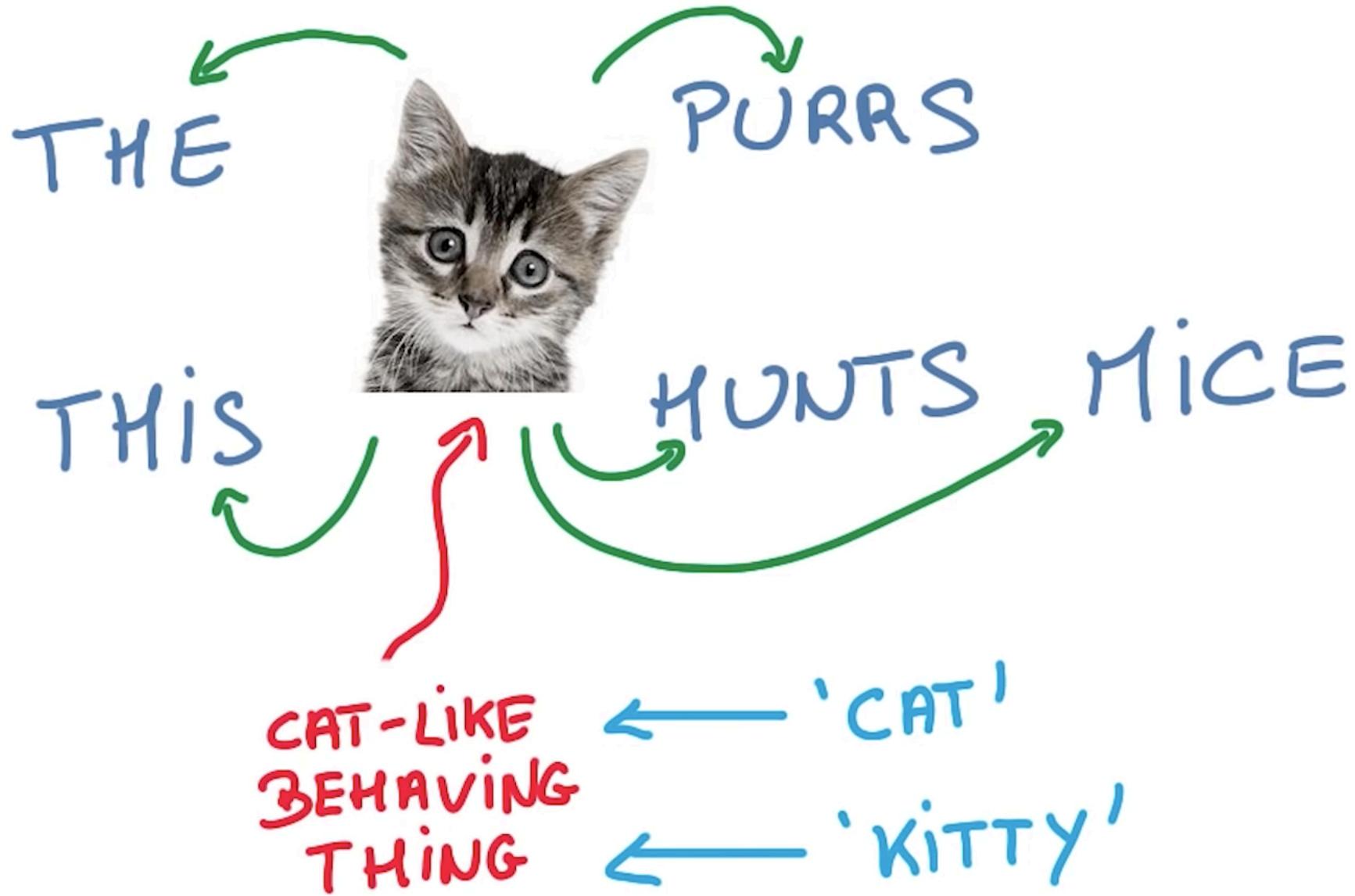
'CAT'



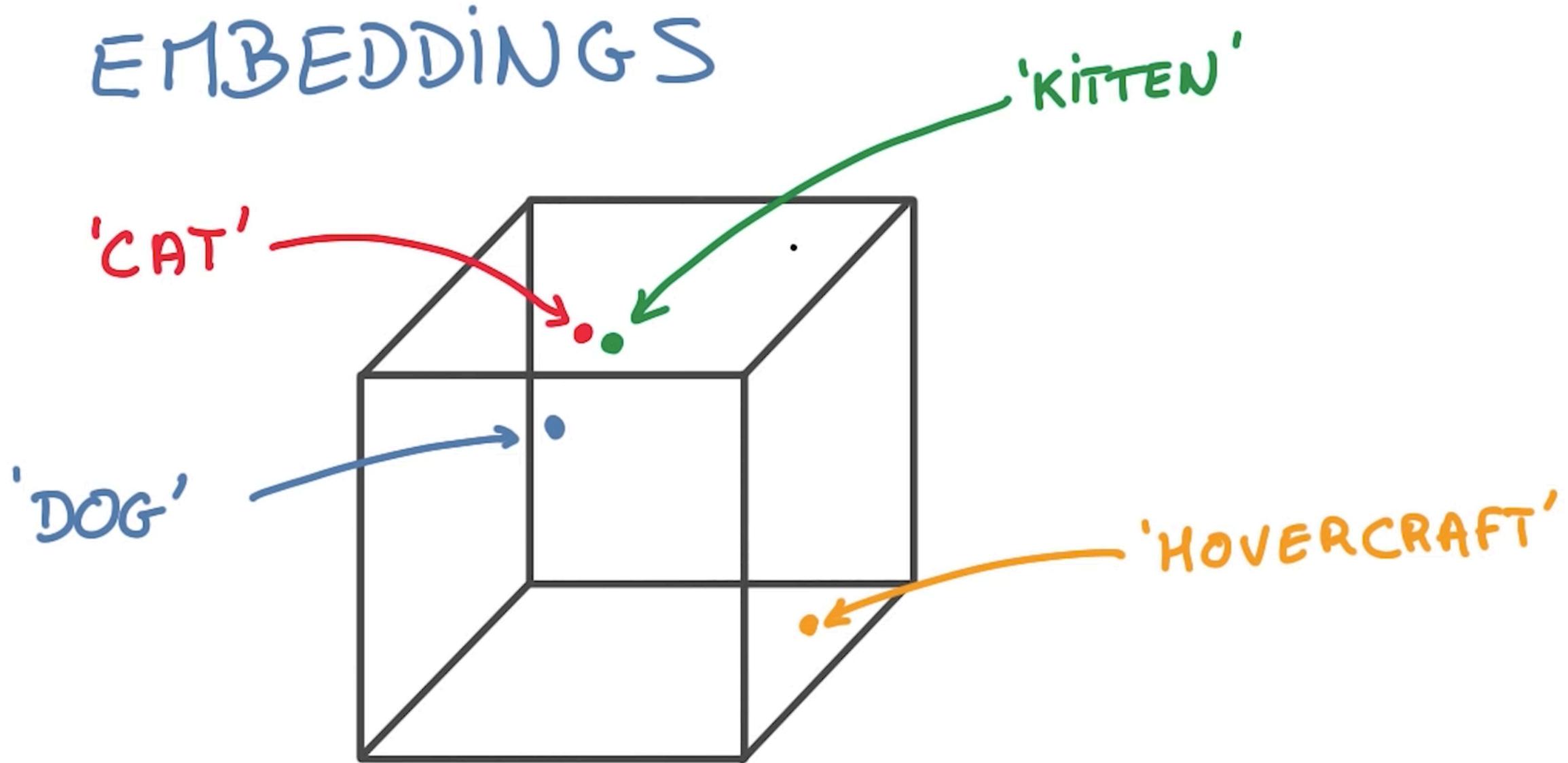
'KITTY'

THE CAT PURRS

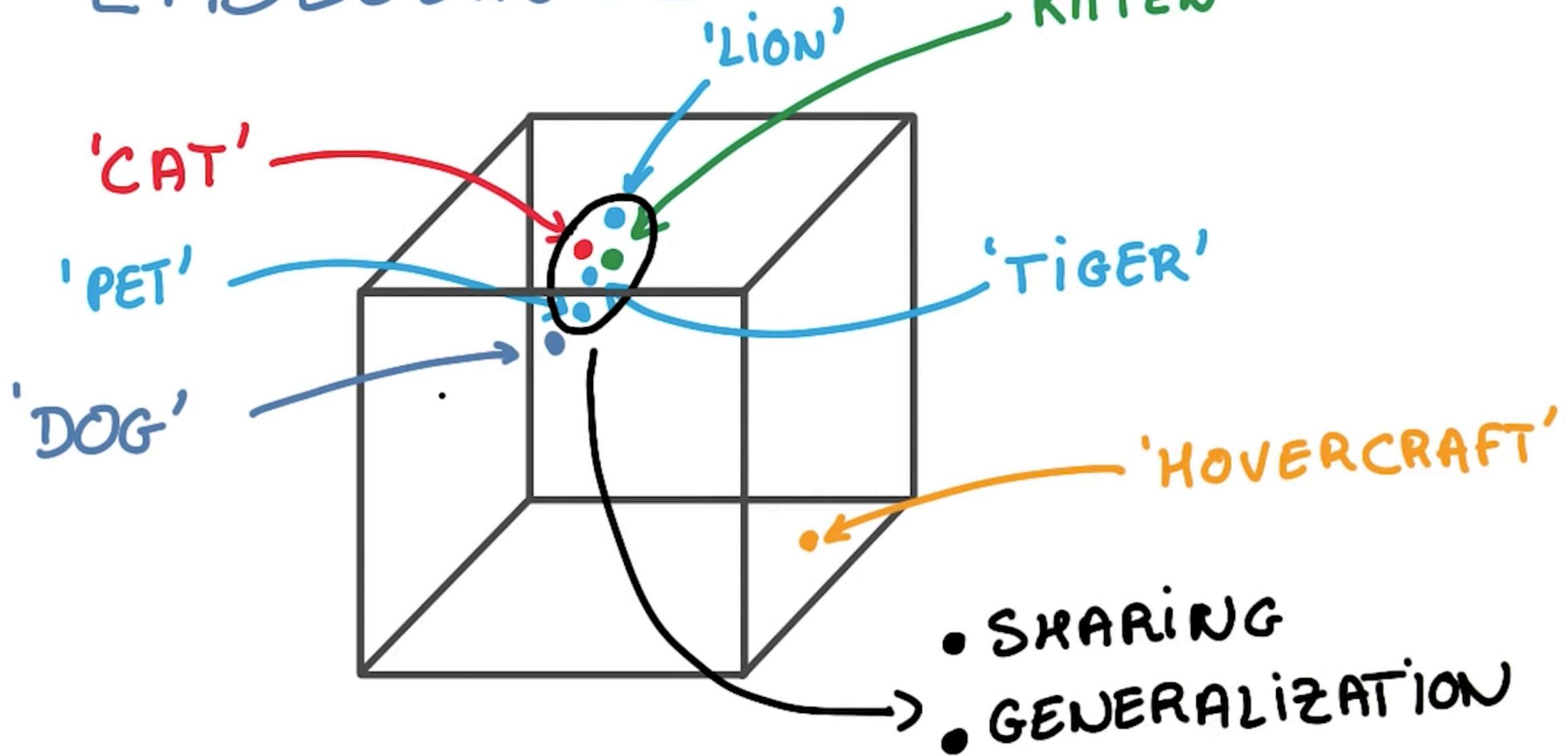
THIS CAT HUNTS MICE



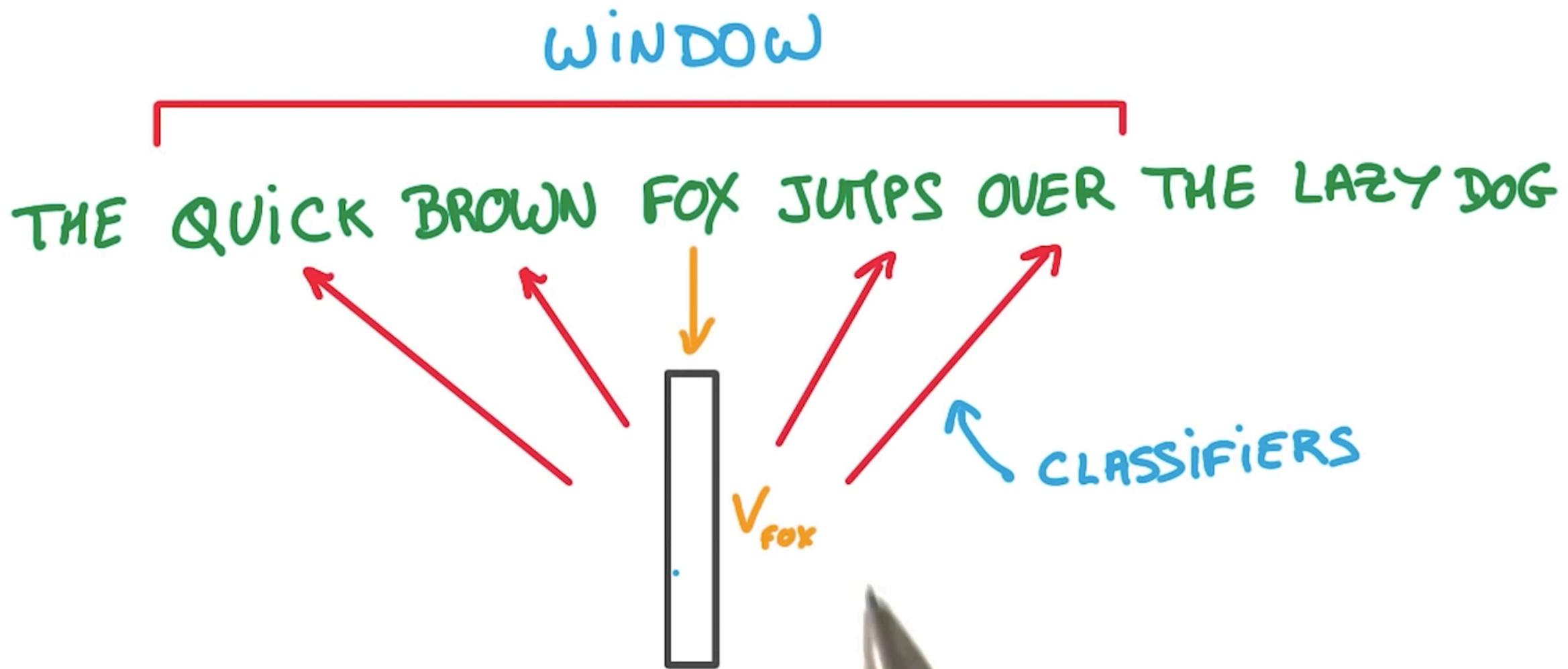
EMBEDDINGS



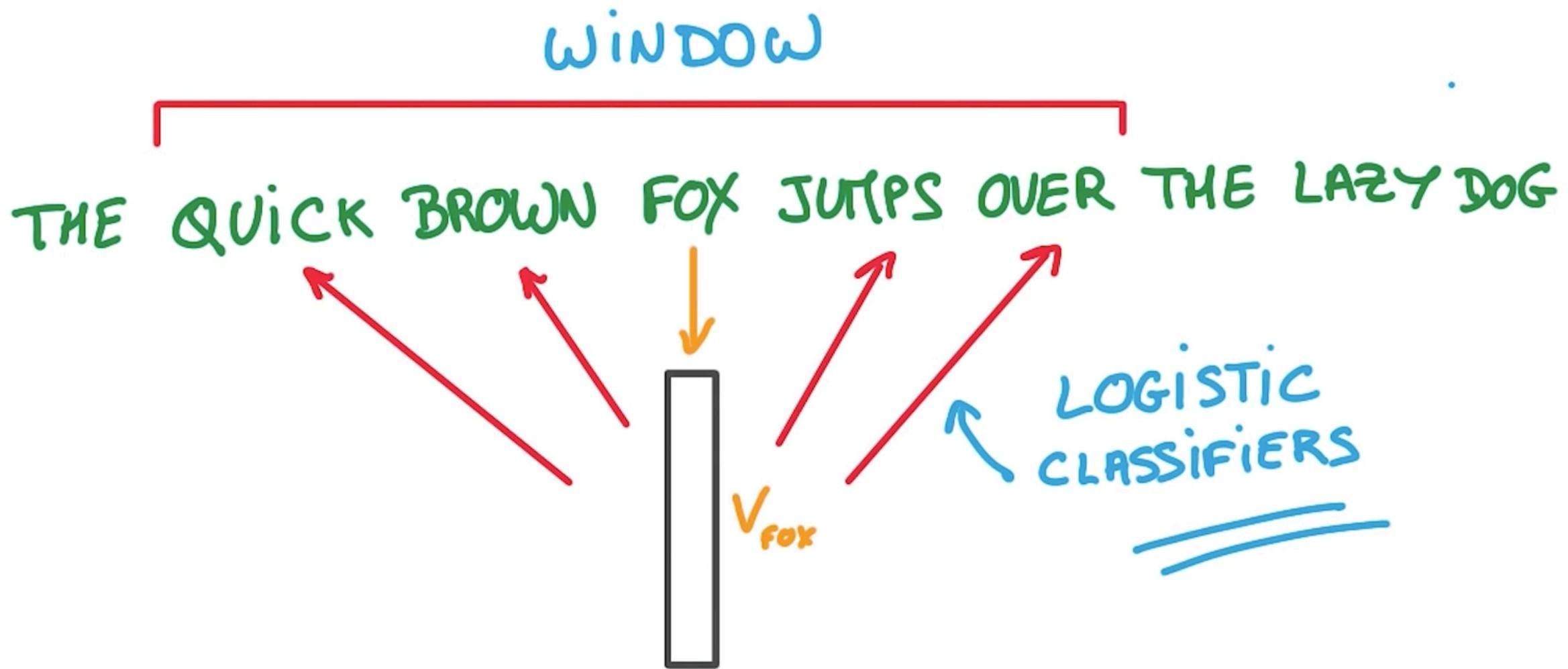
EMBEDDINGS

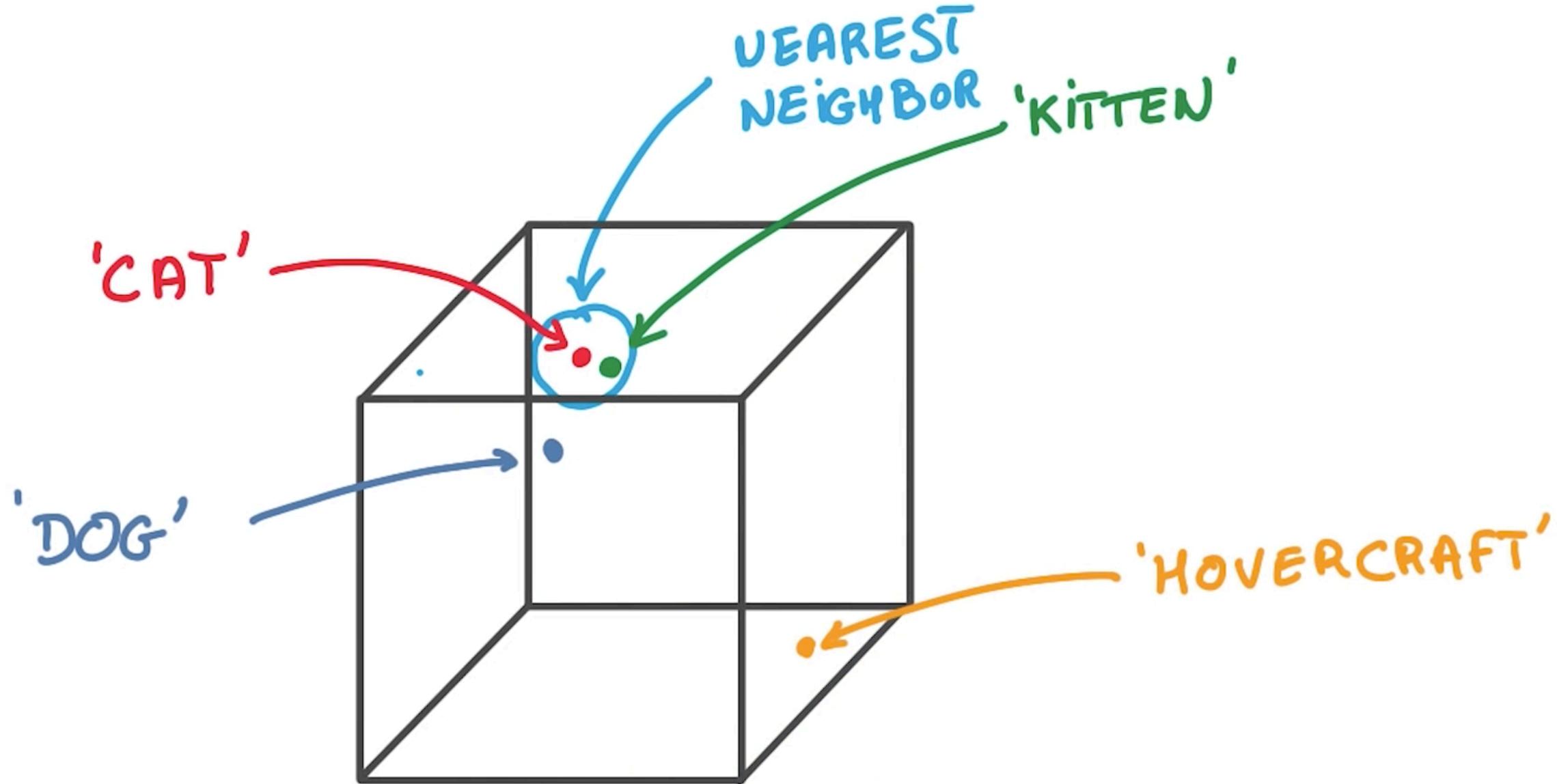


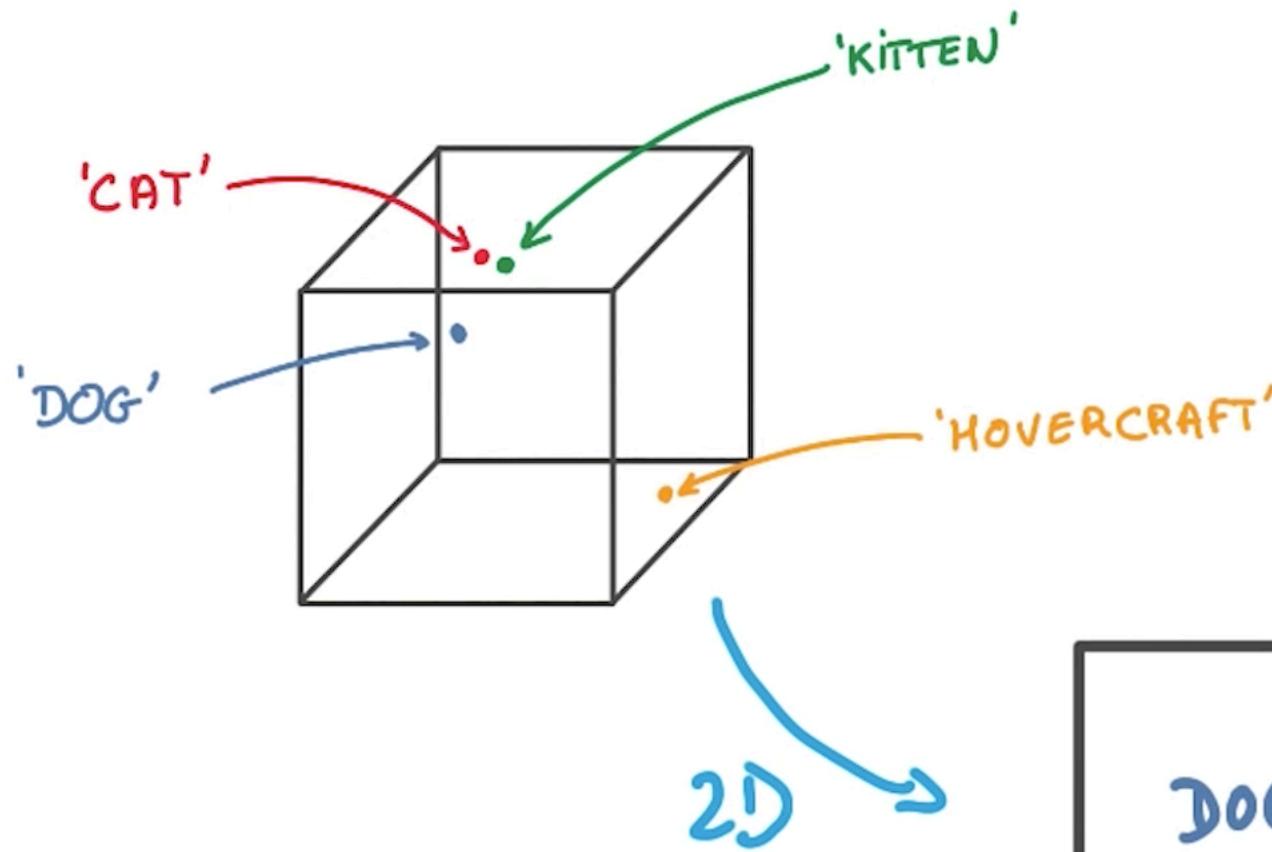
WORD2VEC



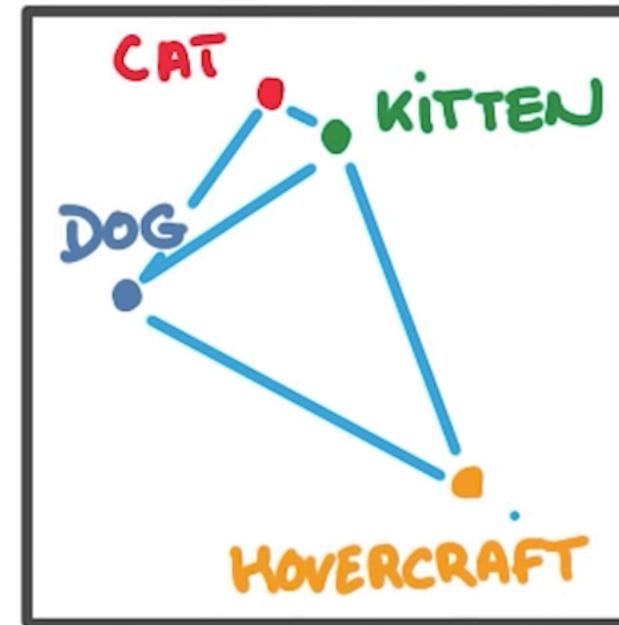
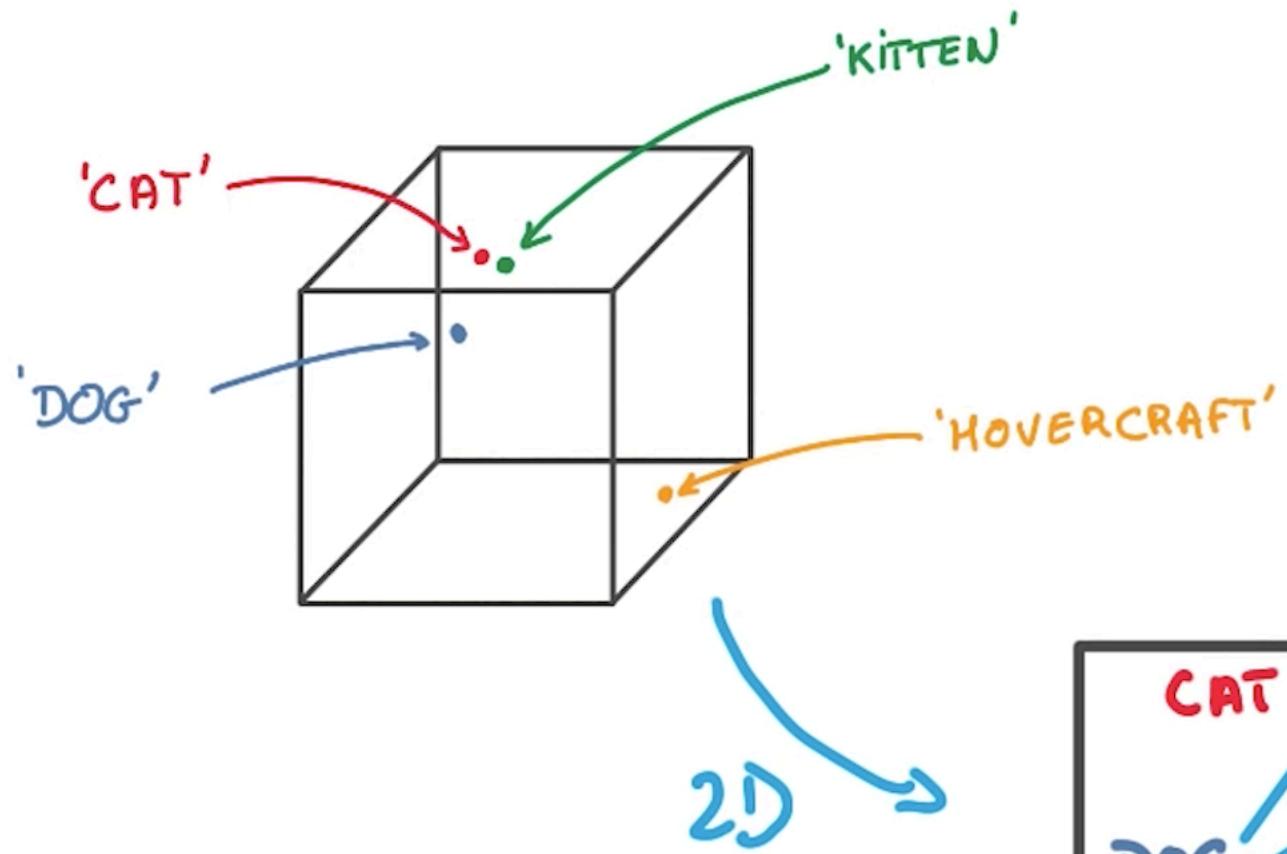
WORD2VEC







~~PCA~~



t-SNE

COMPARING EMBEDDINGS



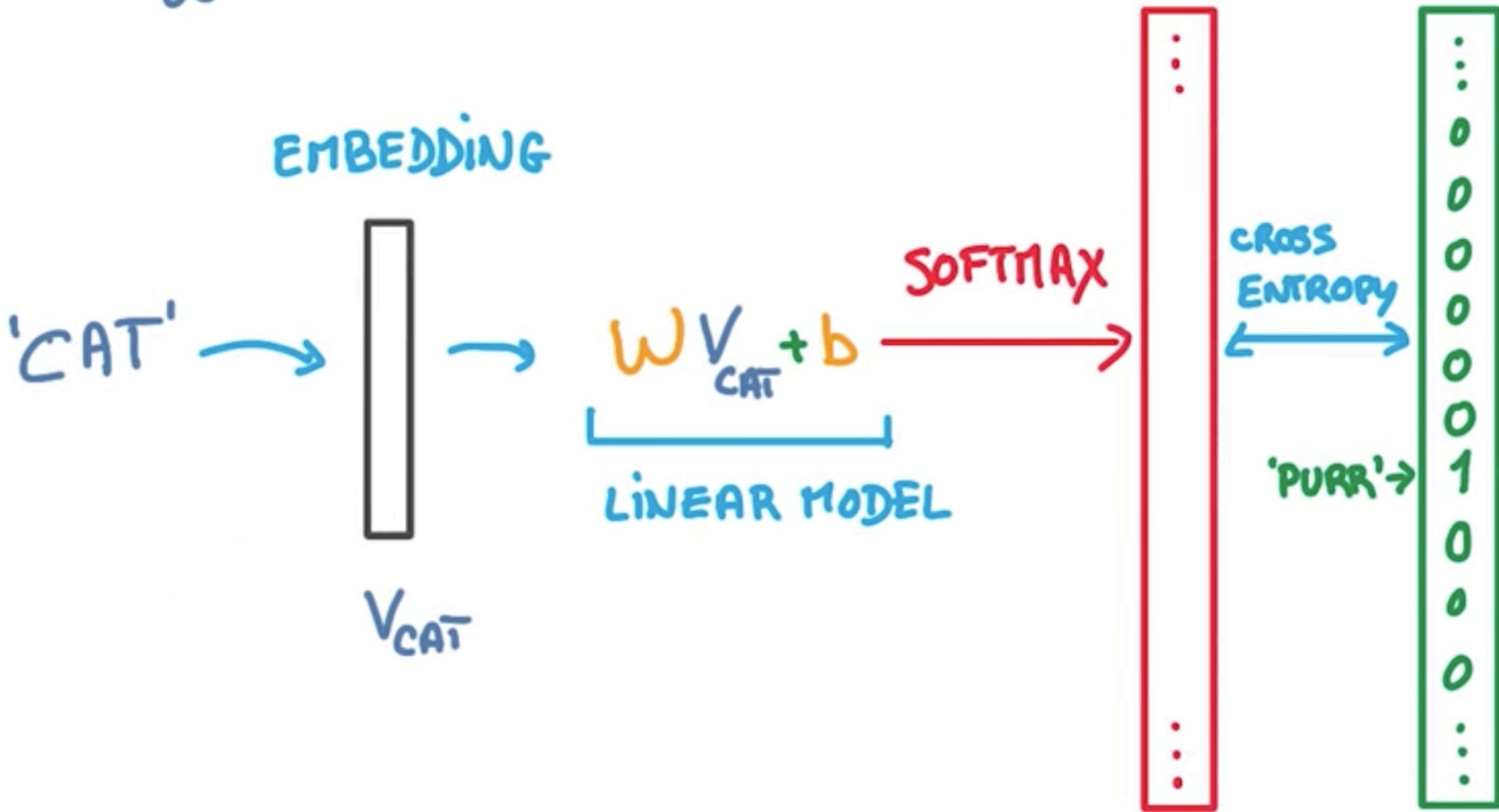
✗ $L_2: \|v_{\text{CAT}} - v_{\text{KITTEN}}\|_2^2$

✓ cosine:

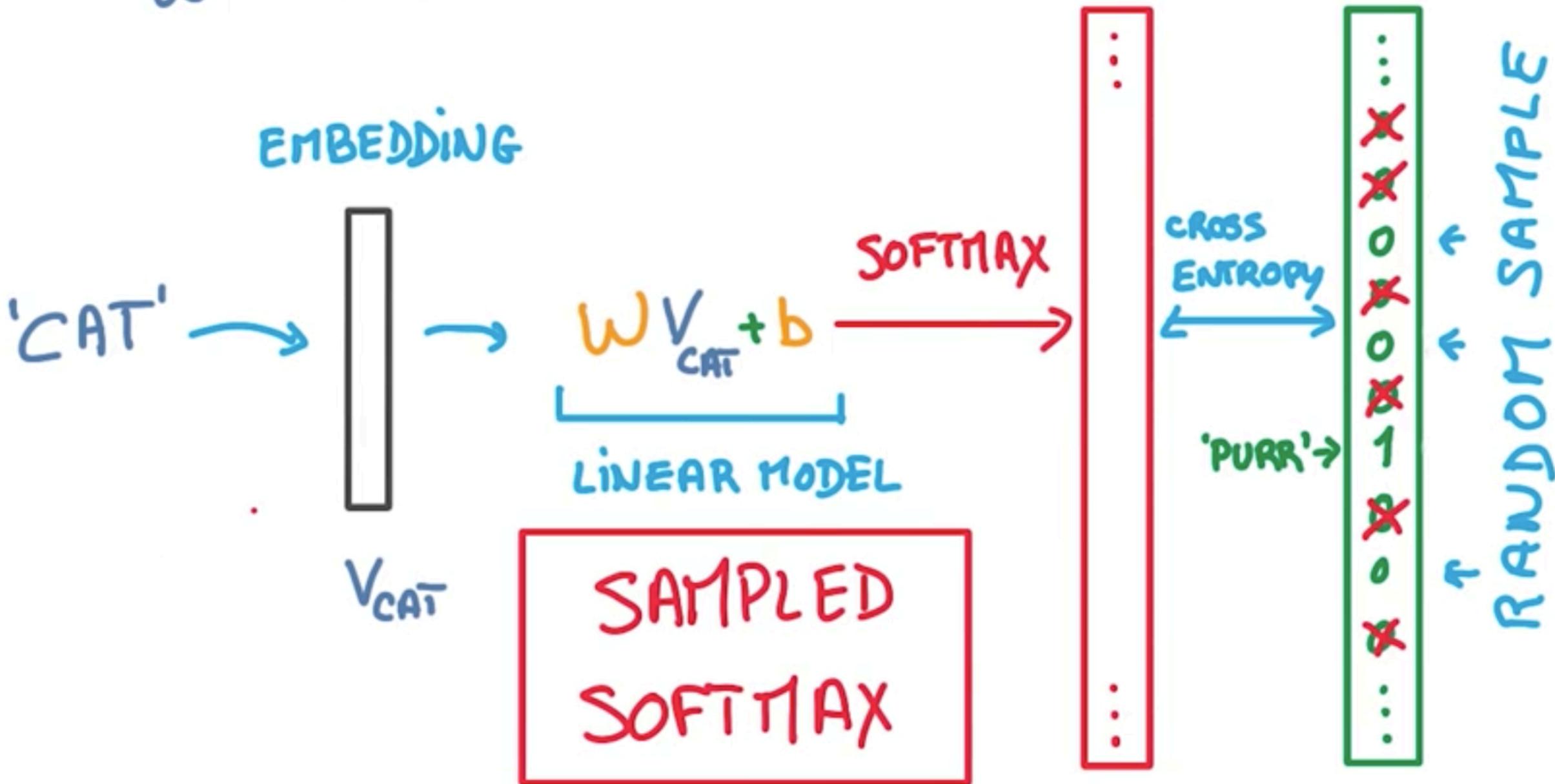
$$\frac{\mathbf{v}_{\text{CAT}} \cdot \mathbf{v}_{\text{KITTEN}}}{\|\mathbf{v}_{\text{CAT}}\| \|\mathbf{v}_{\text{KITTEN}}\|}$$

UNIT NORM

WORD 2 VEC



WORD 2 VEC



WORD ANALOGIES

- PUPPY - DOG + CAT = [DOG
 KITTEN
 WOLF]
- TALLER - TALL + SHORT = [SHORTEST
 TALLEST
 SHORTER]

WORD ANALOGIES

SEMANTIC ANALOGY

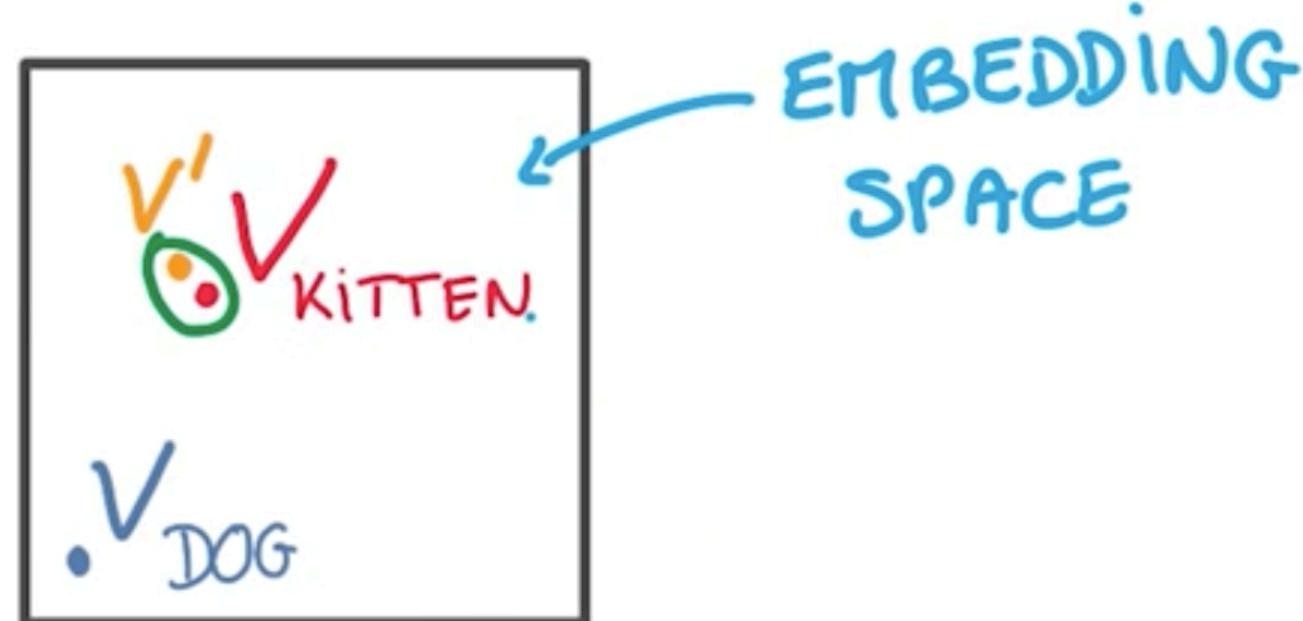
PUPPY → DOG / KITTEN → CAT

SYNTACTIC ANALOGY

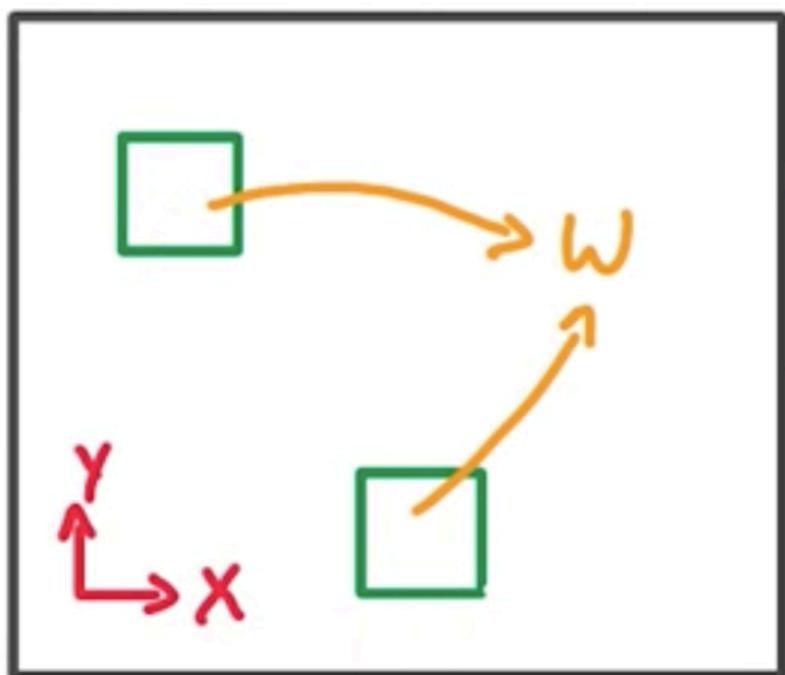
TALLER → TALL / SHORTER → SHORT

WORD ANALOGIES

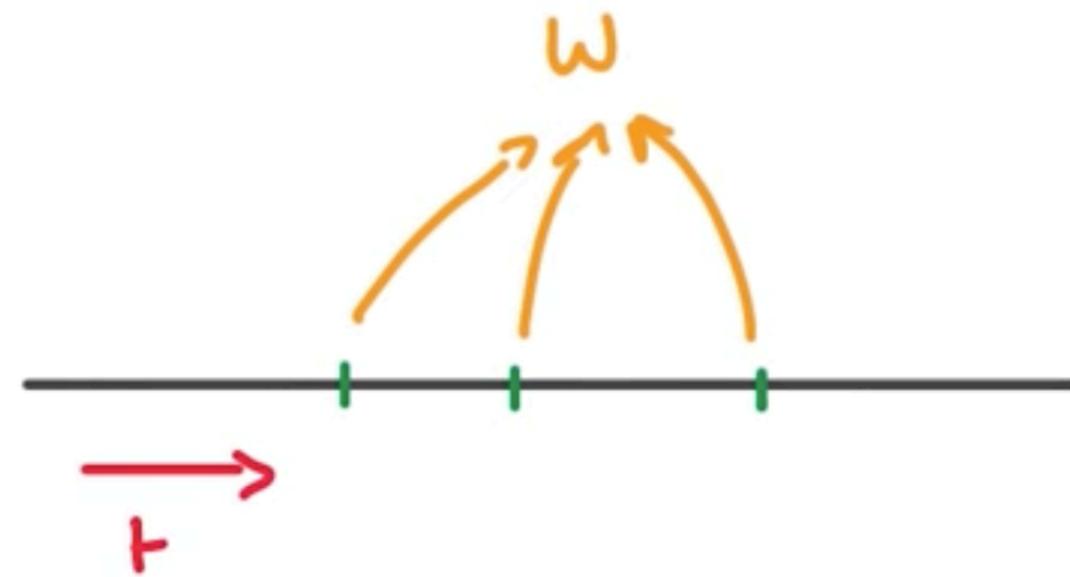
$$V' = V_{\text{PUPPY}} - V_{\text{DOG}} + V_{\text{CAT}}$$



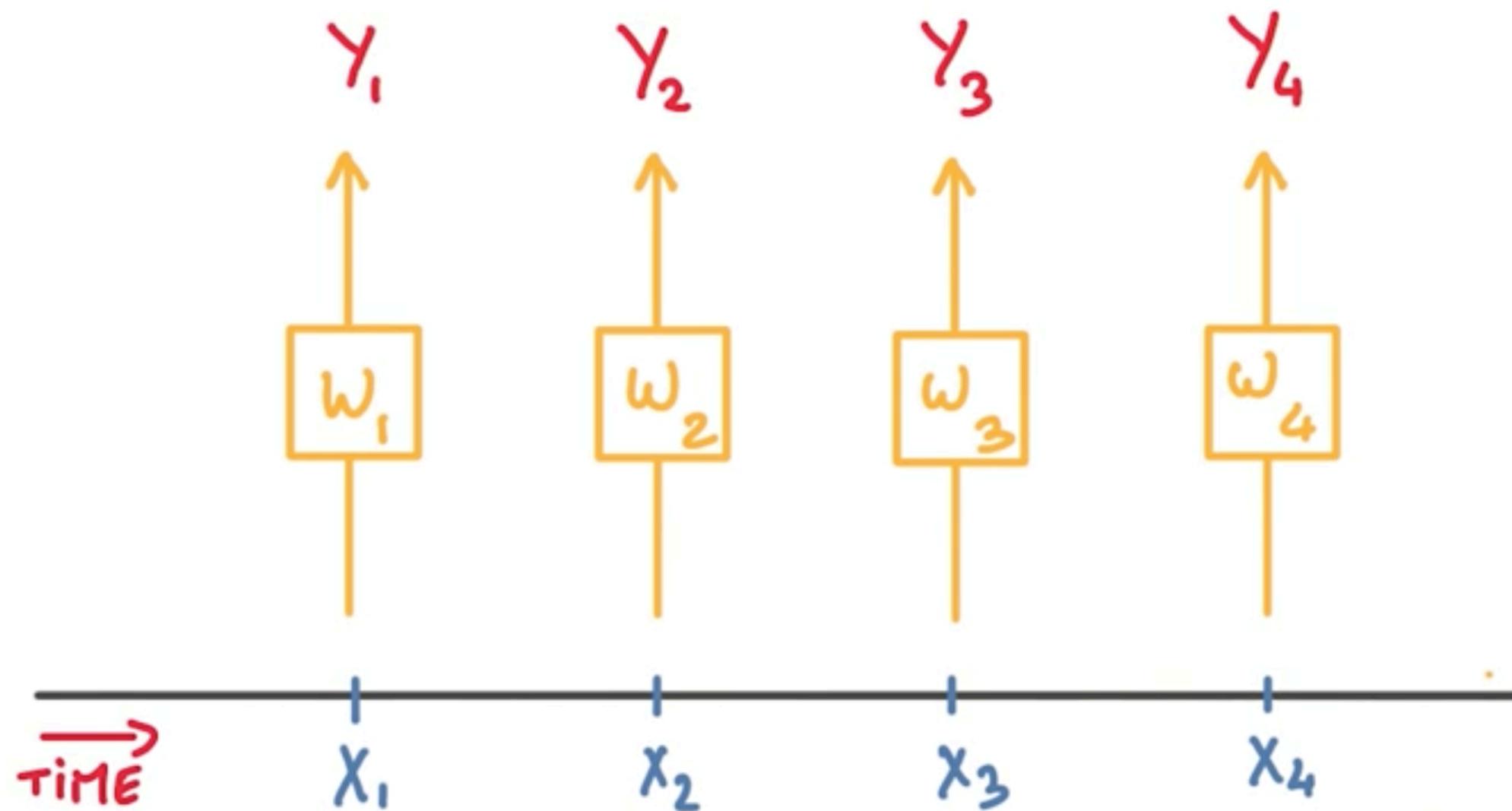
CONVOLUTIONAL NETWORKS



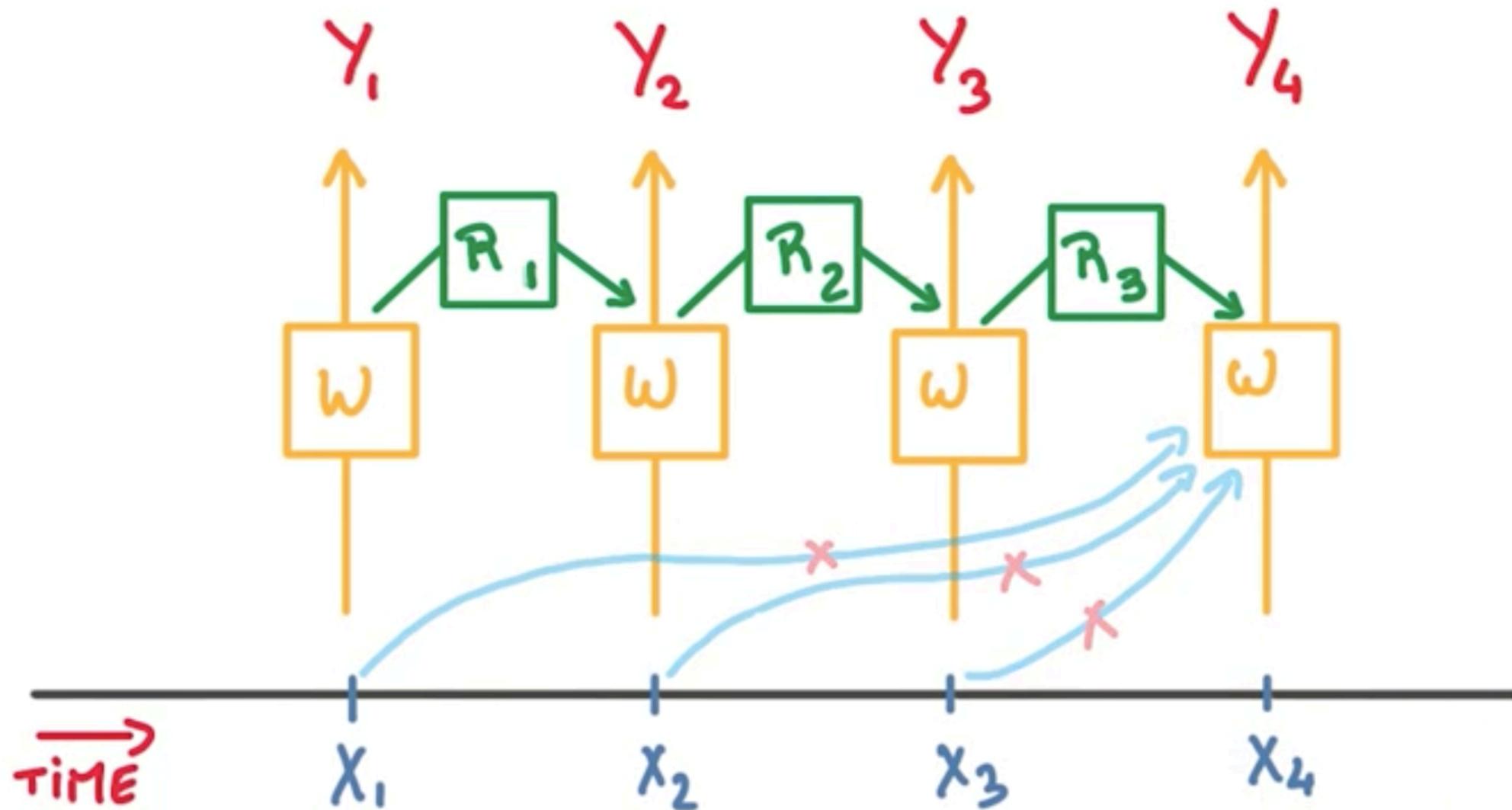
RECURRENT NETWORKS



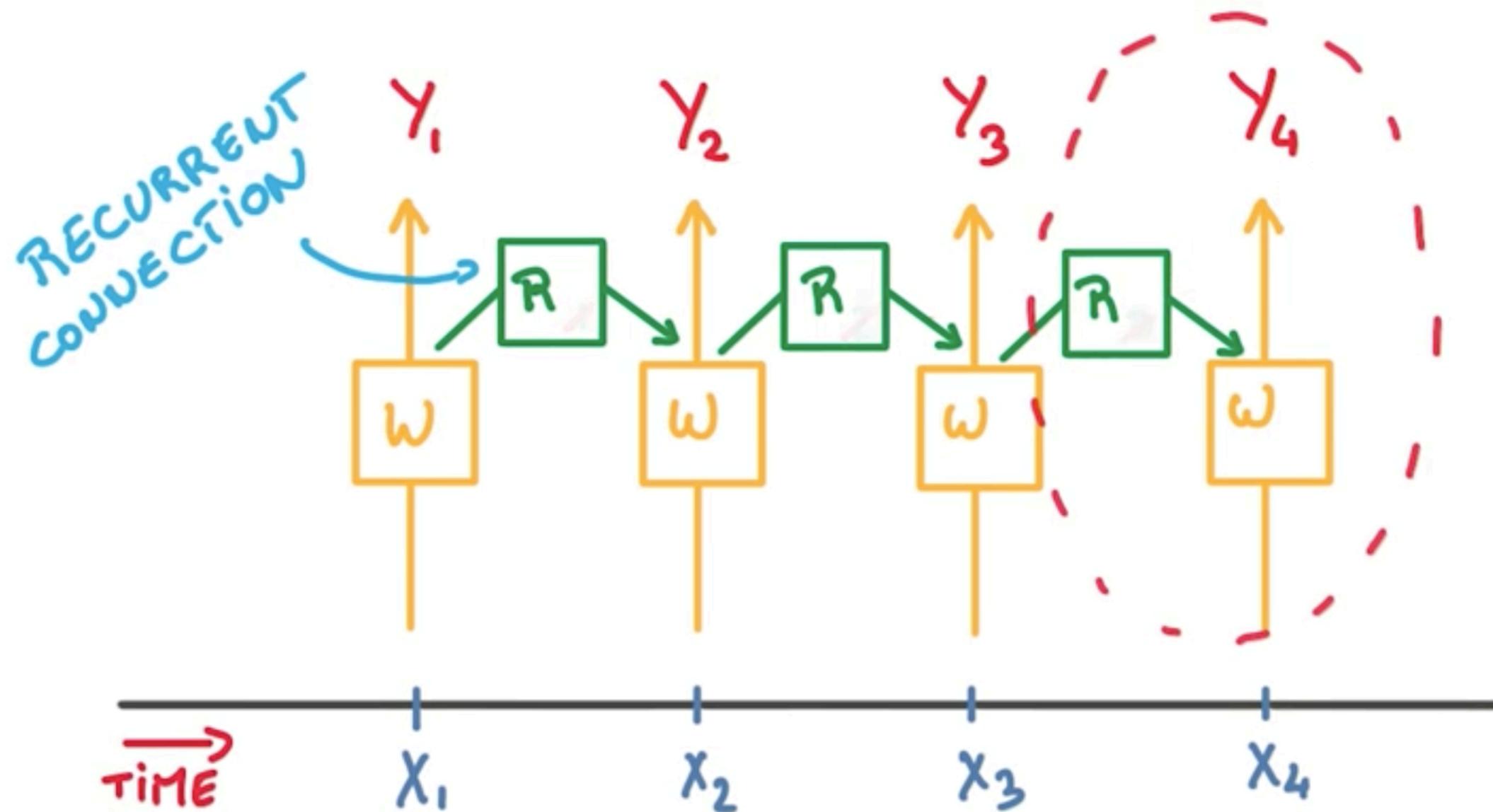
RECURRENT NEURAL NETWORKS



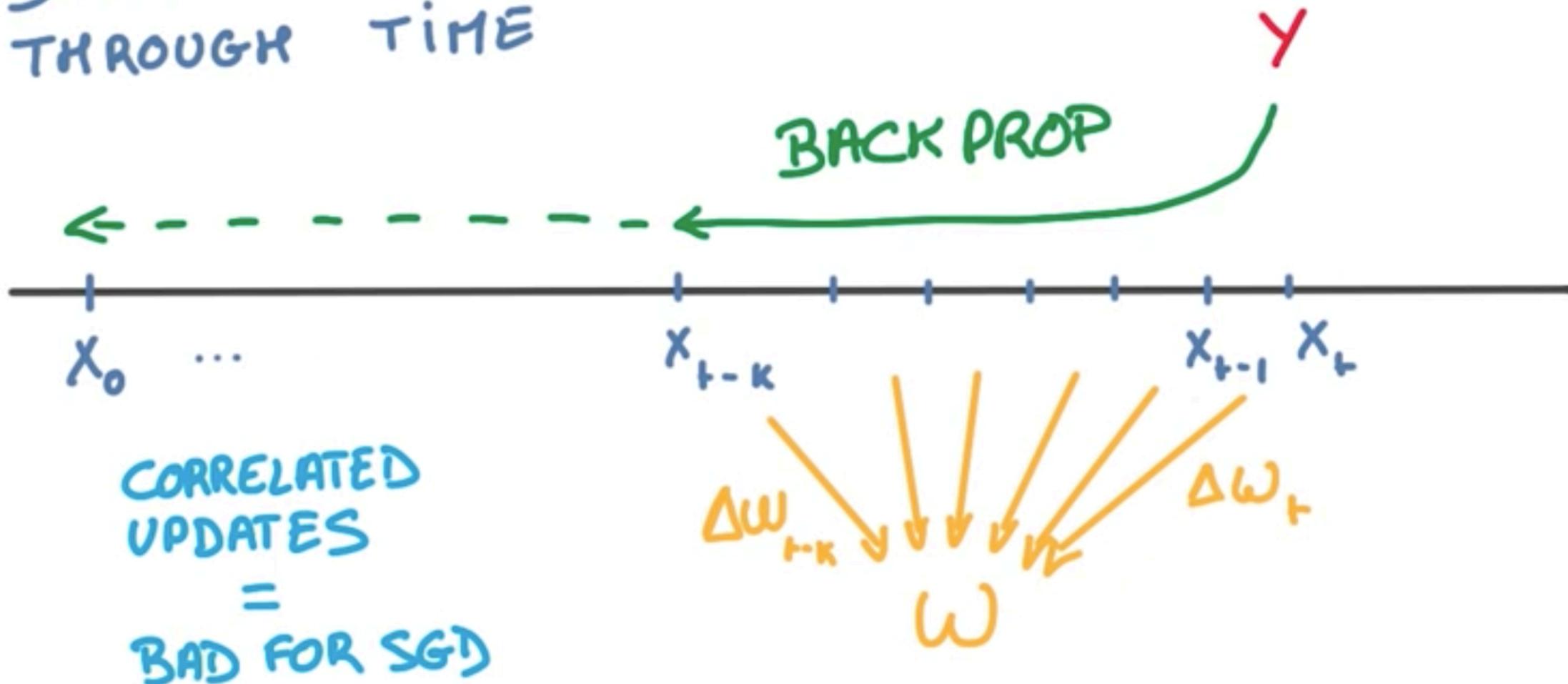
RECURRENT NEURAL NETWORKS



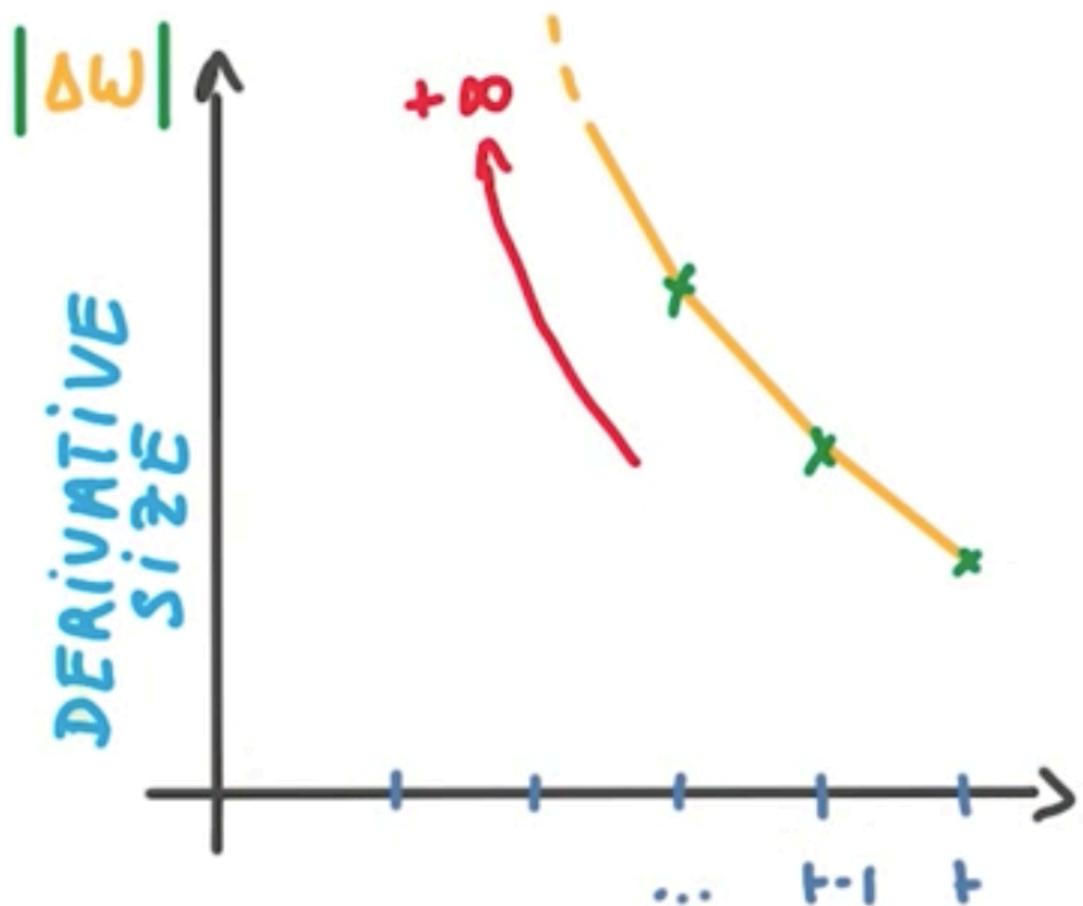
RECURRENT NEURAL NETWORKS



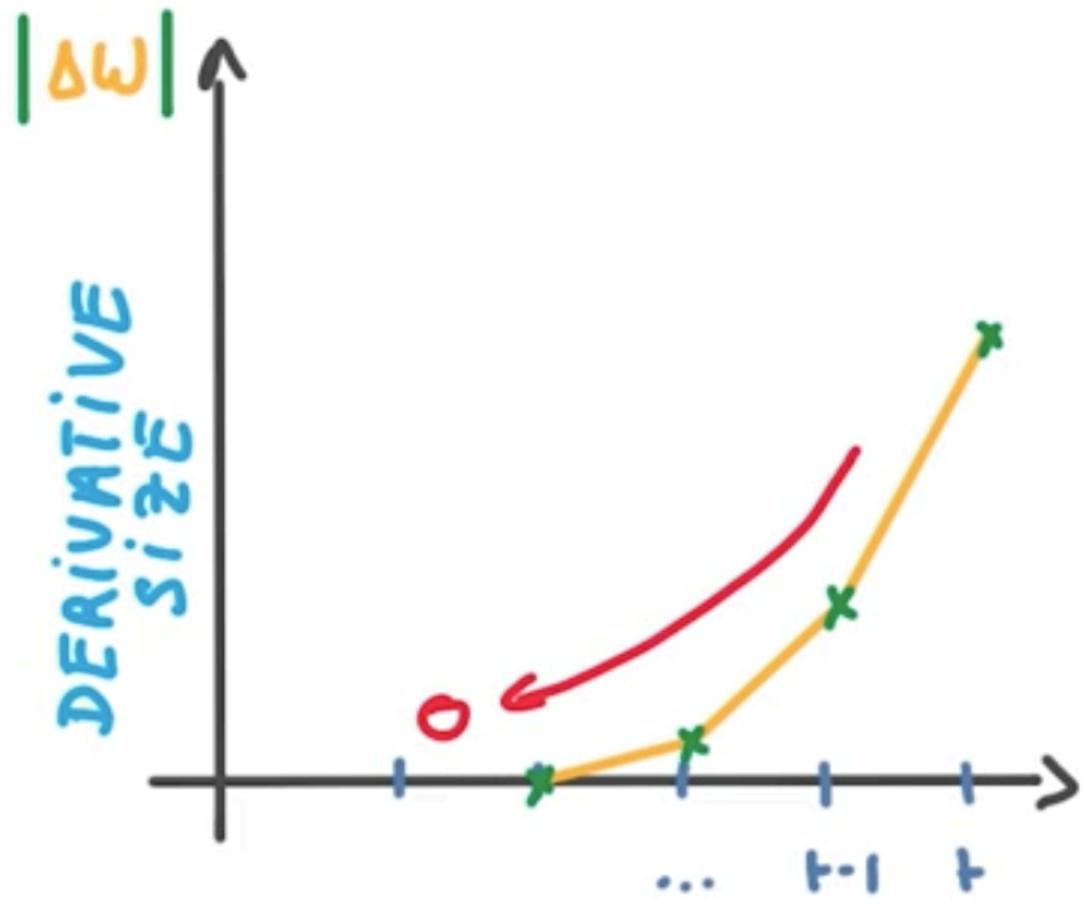
BACKPROPAGATION THROUGH TIME



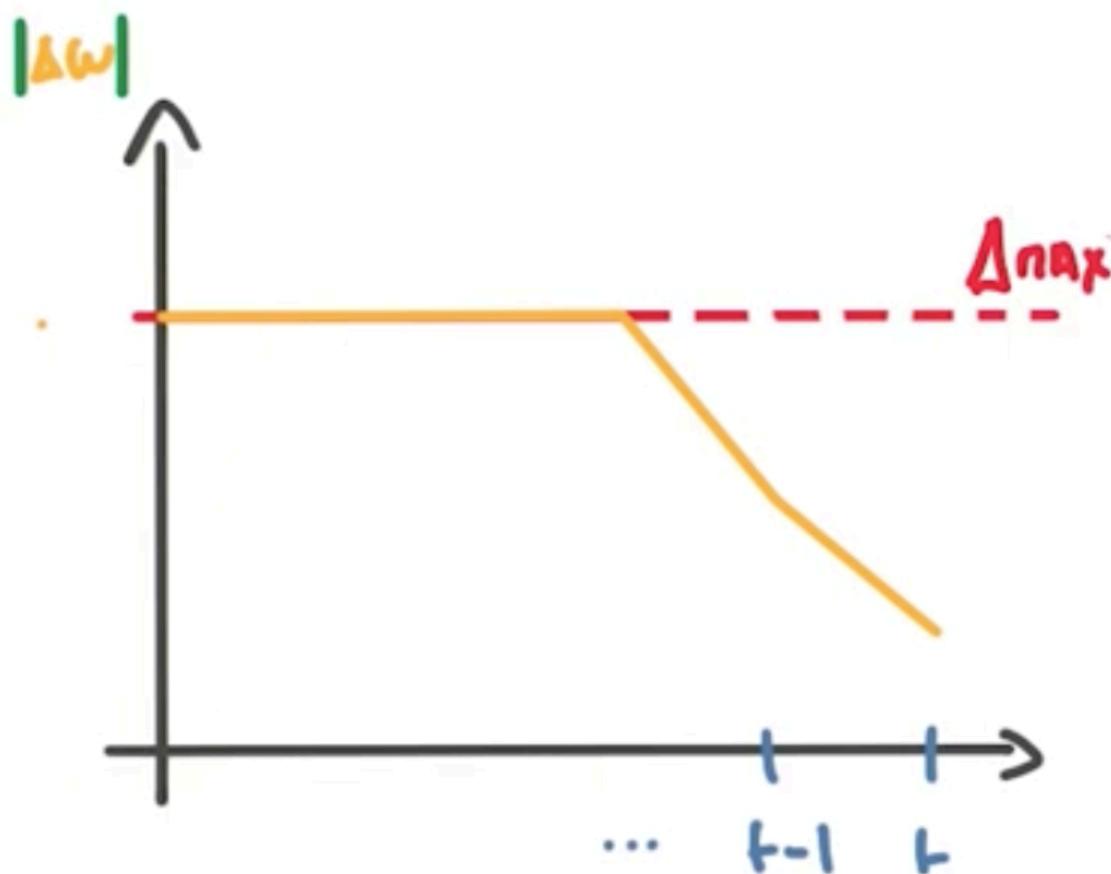
EXPLODING GRADIENT



VANISHING GRADIENT



EXPLODING GRADIENTS: GRADIENT CLIPPING



$$\Delta\omega \leftarrow \Delta\omega \frac{\Delta_{\max}}{\max(|\Delta\omega|, \Delta_{\max})}$$

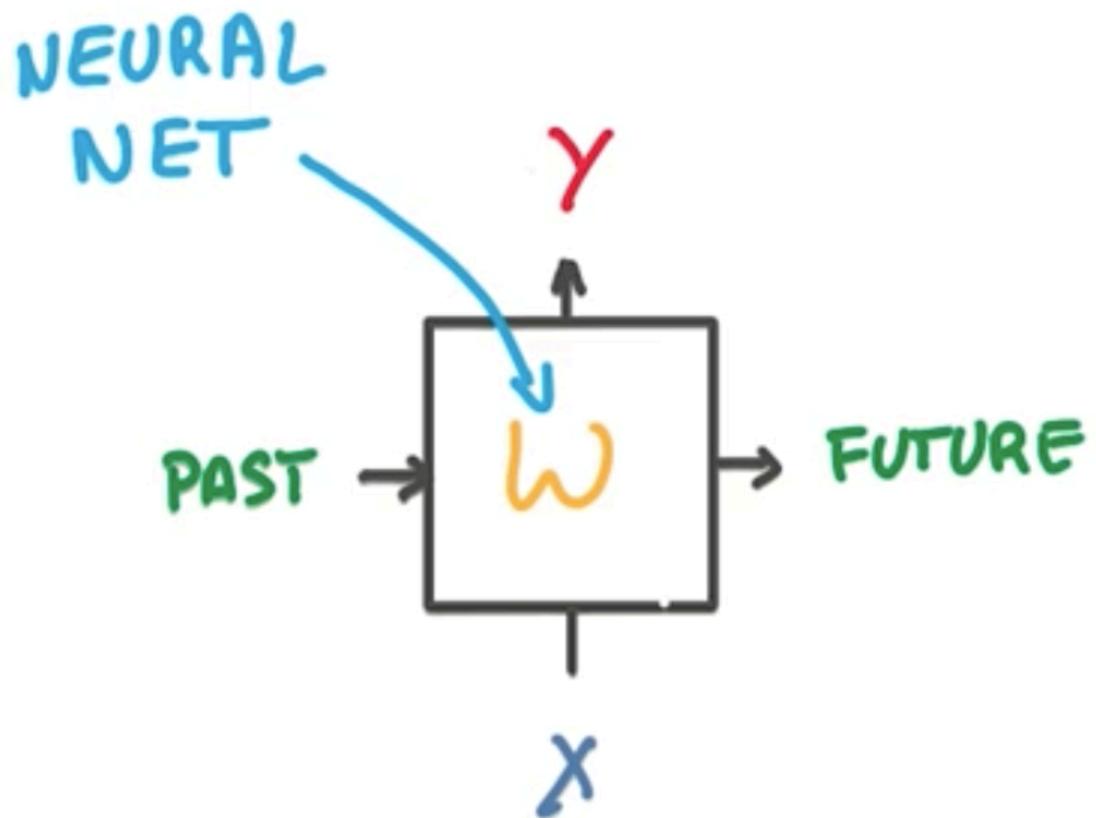
VANISHING GRADIENTS

'MEMORY LOSS' IN RNNs

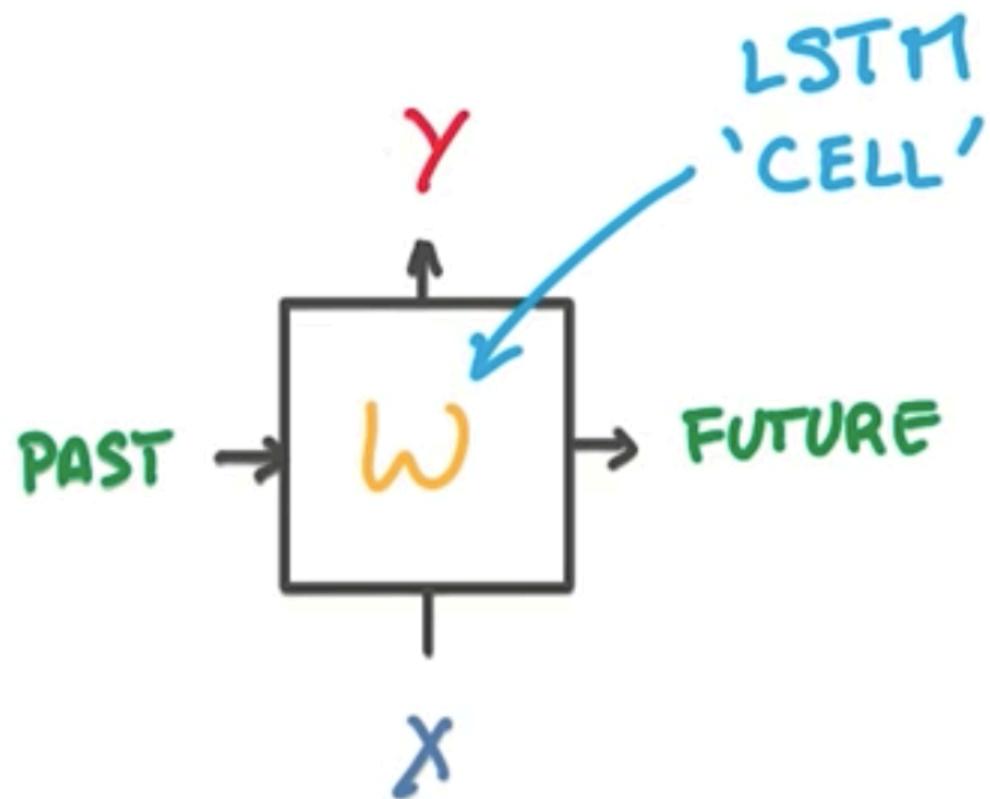
SO A ZSL LUCKY BROWN FOX JUMPS ...

The diagram shows a sequence of words: SO A ZSL LUCKY BROWN FOX JUMPS ... Above the sequence, there are two question marks with dashed arrows pointing upwards, representing the flow of gradients. A green curved arrow originates from the word 'FOX' and points to the first question mark, symbolizing that the gradient information is lost as it propagates through the sequence.

LSTM (LONG SHORT-TERM MEMORY)

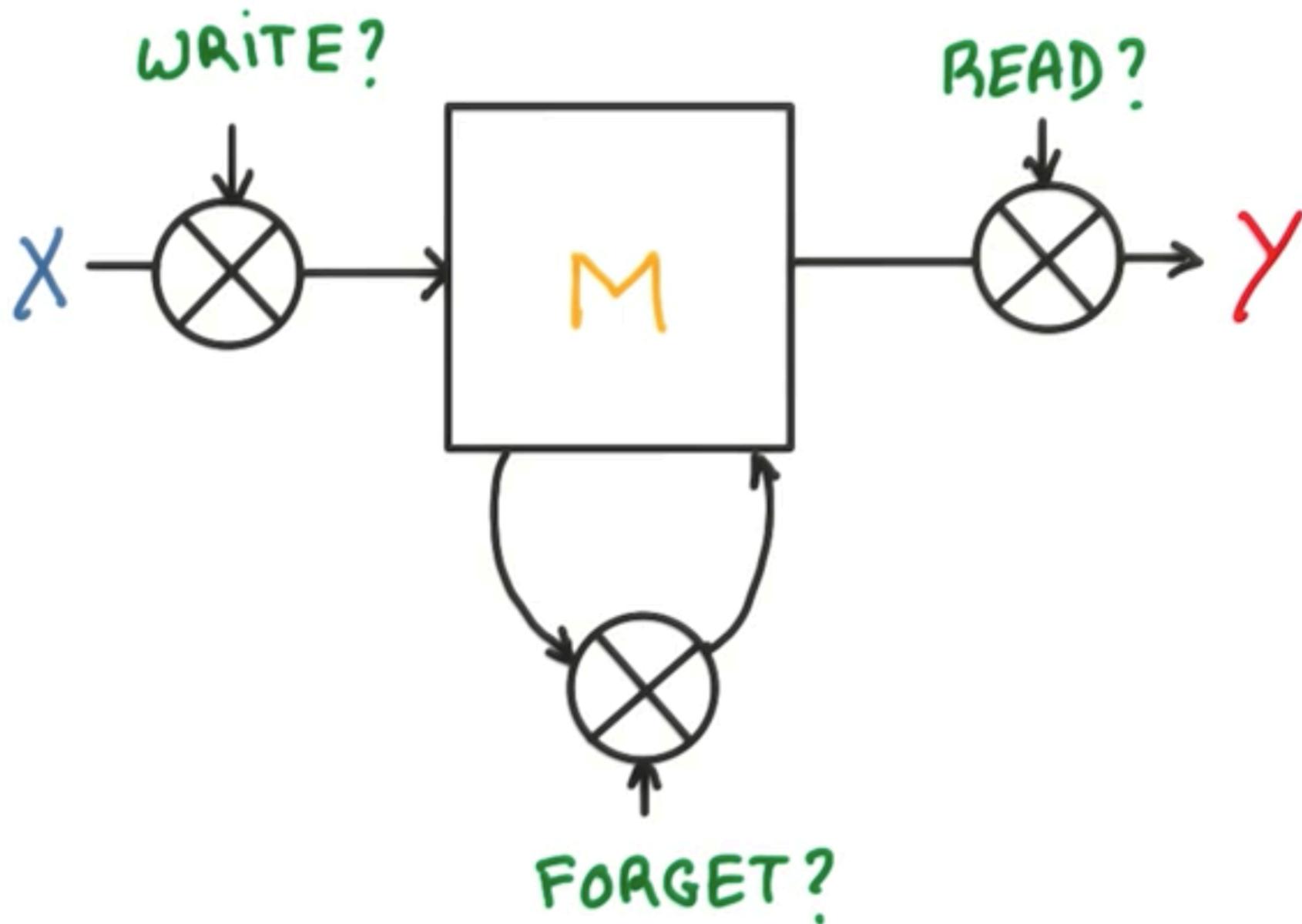


LSTM (LONG SHORT-TERM MEMORY)

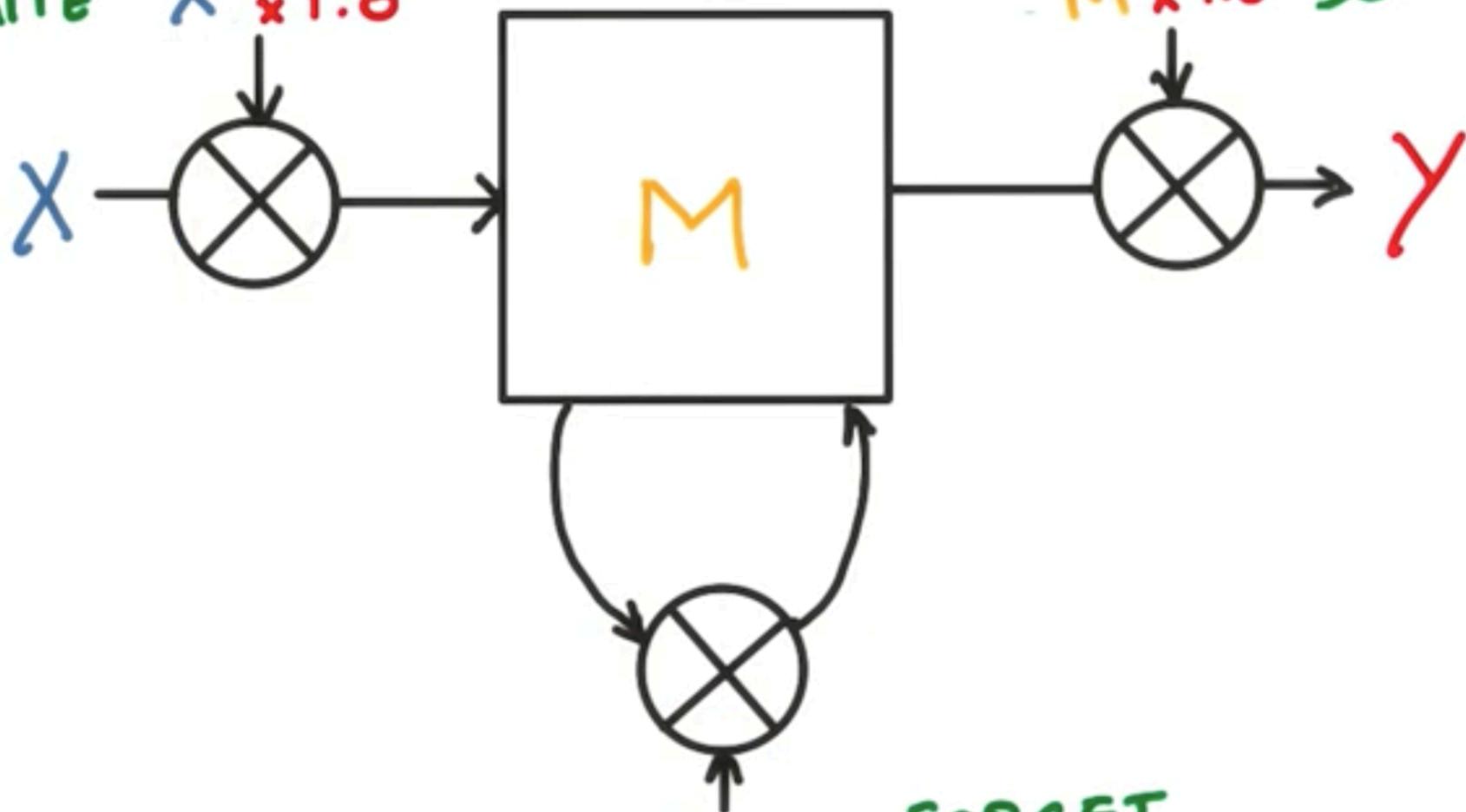


A SIMPLE MODEL OF MEMORY





DON'T WRITE $X \times 0.0$
DO WRITE $X \times 1.0$

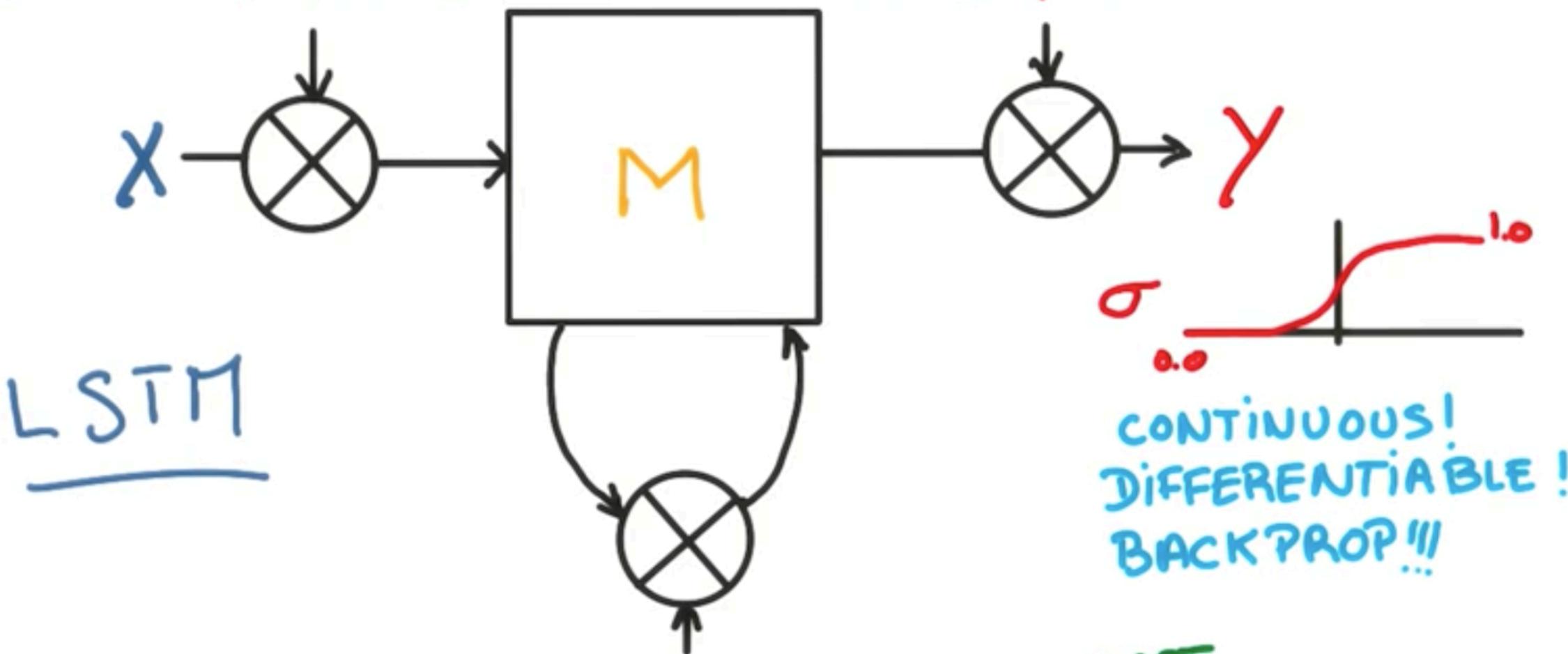


M $\times 0.0$ DON'T READ
M $\times 1.0$ DO READ

M $\times 0.0$
M $\times 1.0$
FORGET
DON'T FORGET

MAYBE WRITE $X \times \sigma \in [0.0, 1.0]$

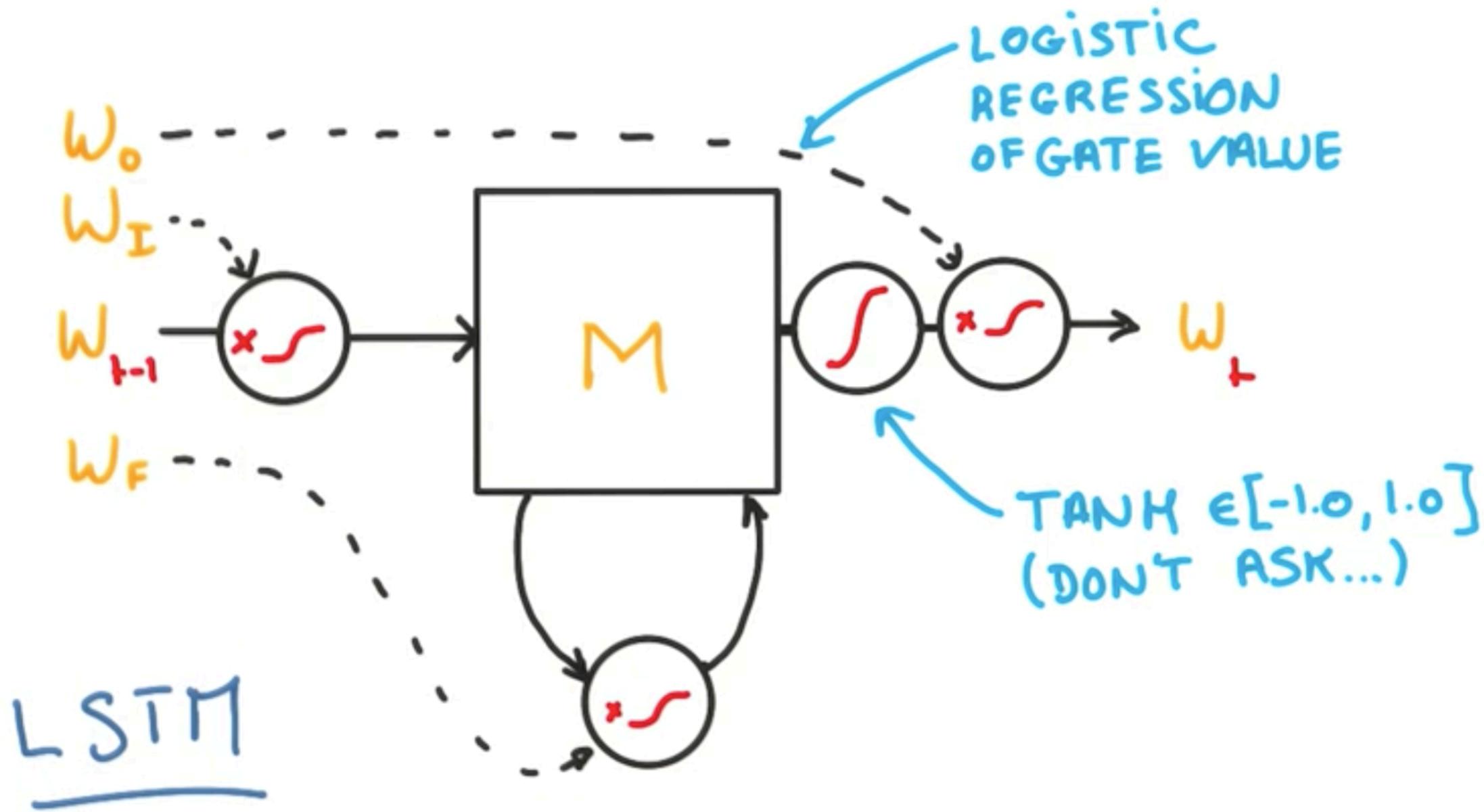
$\Pi \times \sigma \in [0.0, 1.0]$ MAYBE READ



LSTM

CONTINUOUS!
DIFFERENTIABLE!
BACKPROP!!!

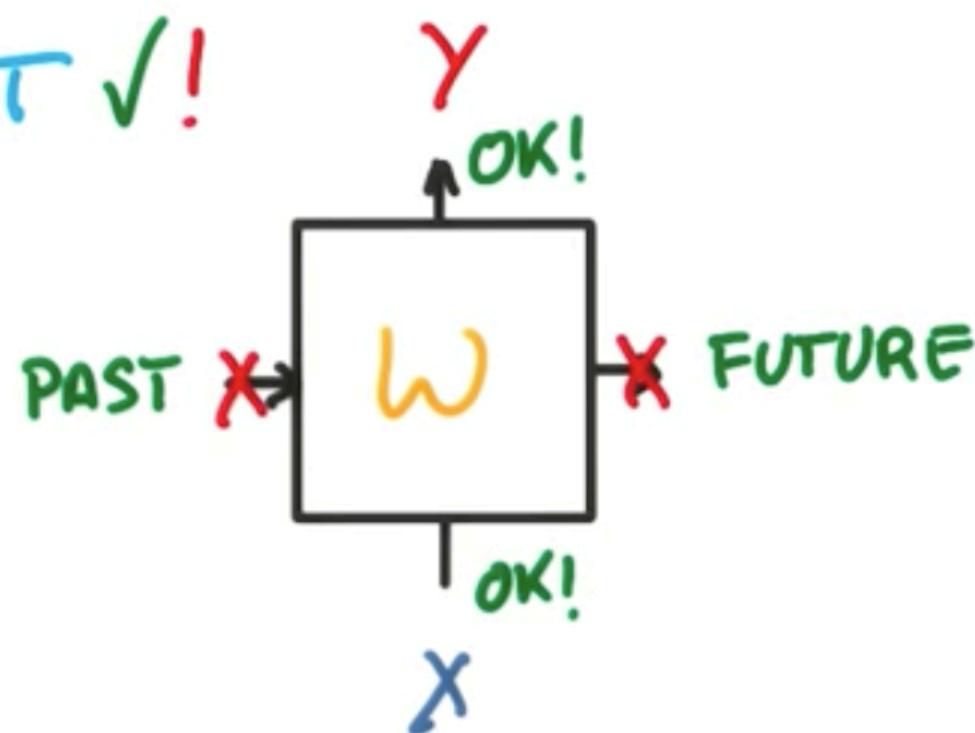
$\Pi \times \sigma \in [0.0, 1.0]$ MAYBE FORGET



LSTM REGULARIZATION

$L_2 \checkmark$

DROPOUT $\checkmark!$



THE QUICK BROWN FOX...

THE QUICK BRO...

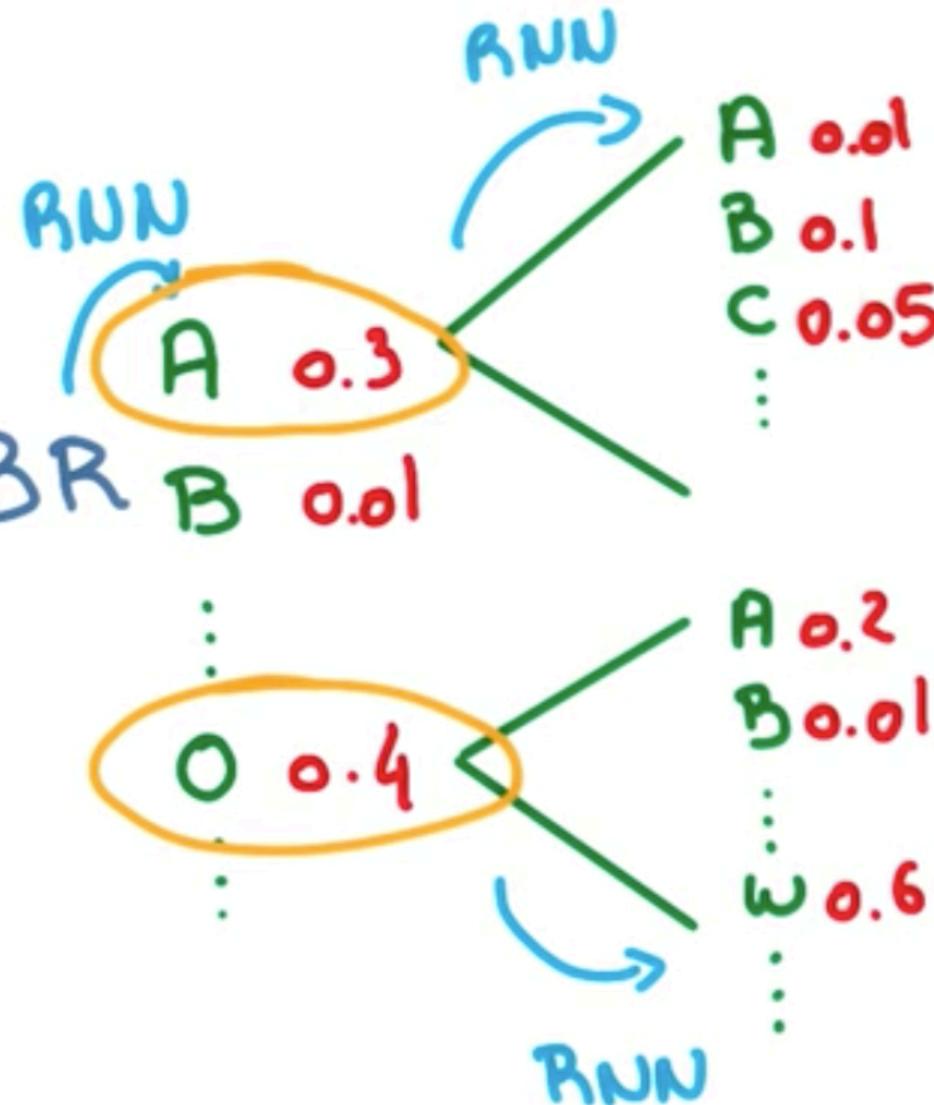




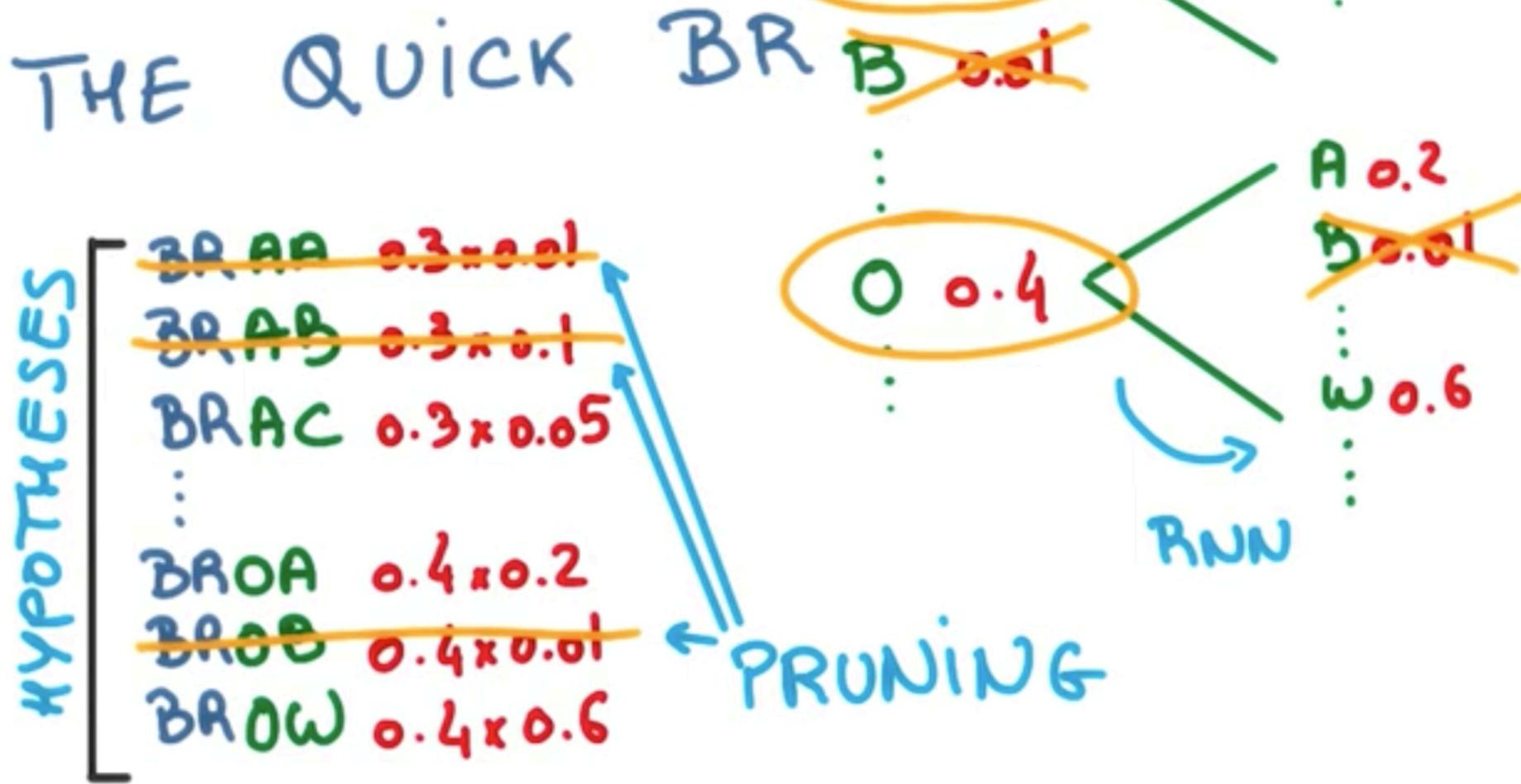
THE QUICK BR

HYPOTHESES

- BR AA 0.3×0.01
- BR AB 0.3×0.1
- BR AC 0.3×0.05
- ⋮
- BR OA 0.4×0.2
- BR OB 0.4×0.01
- BR OW 0.4×0.6**



BEAM SEARCH

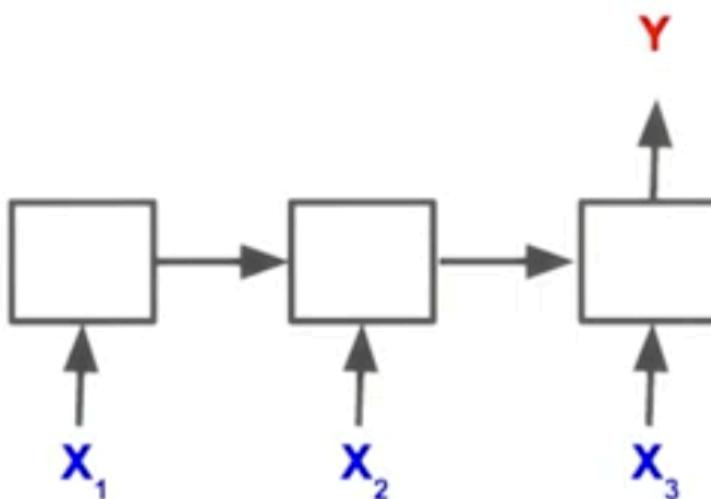


VARIABLE-LENGTH
SEQUENCE IN



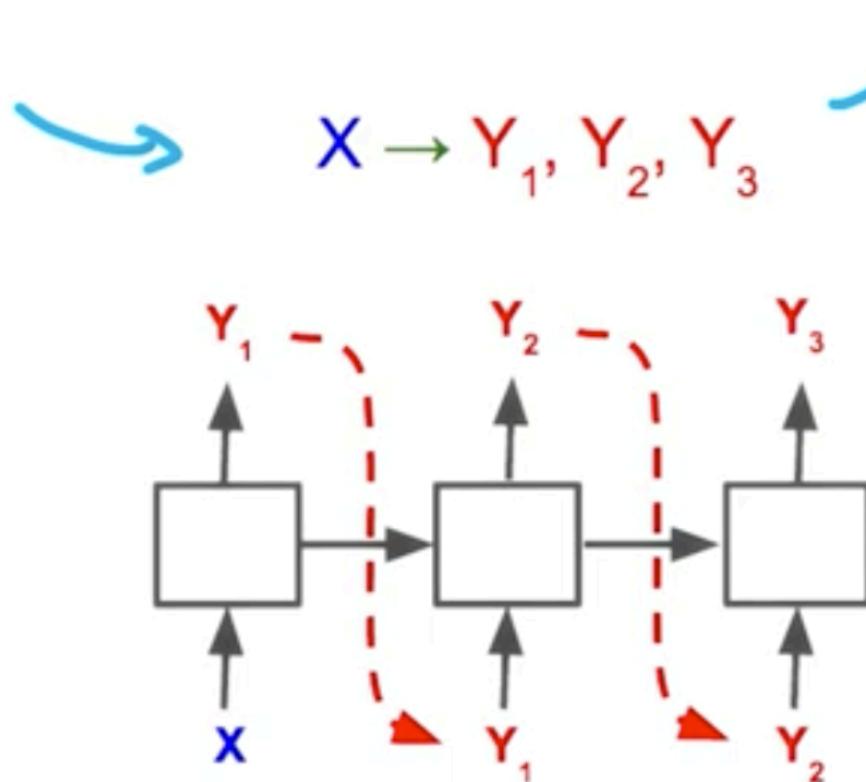
$X_1, X_2, X_3 \rightarrow Y$

FIXED-LENGTH
VECTOR OUT



FIXED-LENGTH
VECTOR IN

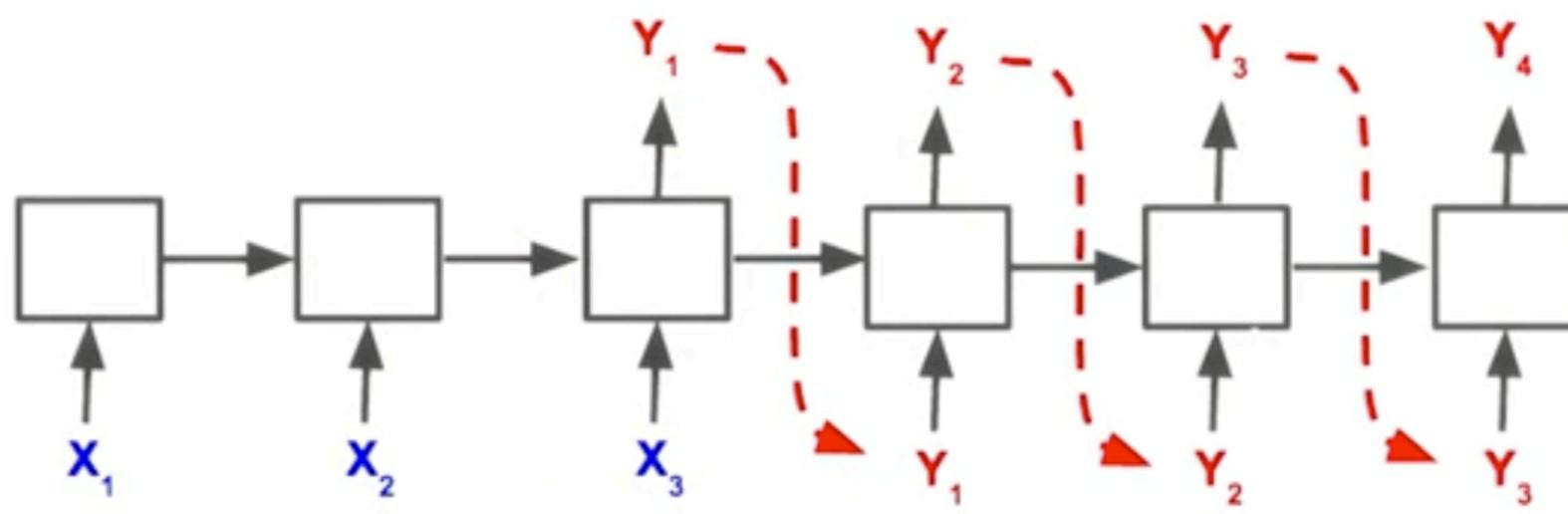
VARIABLE-LENGTH
SEQUENCE OUT



SEQUENCE IN

SEQUENCE OUT

$$X_1, X_2, X_3 \rightarrow Y_1, Y_2, Y_3, Y_4$$



$$X_1, X_2, X_3 \rightarrow Y_1, Y_2, Y_3, Y_4$$

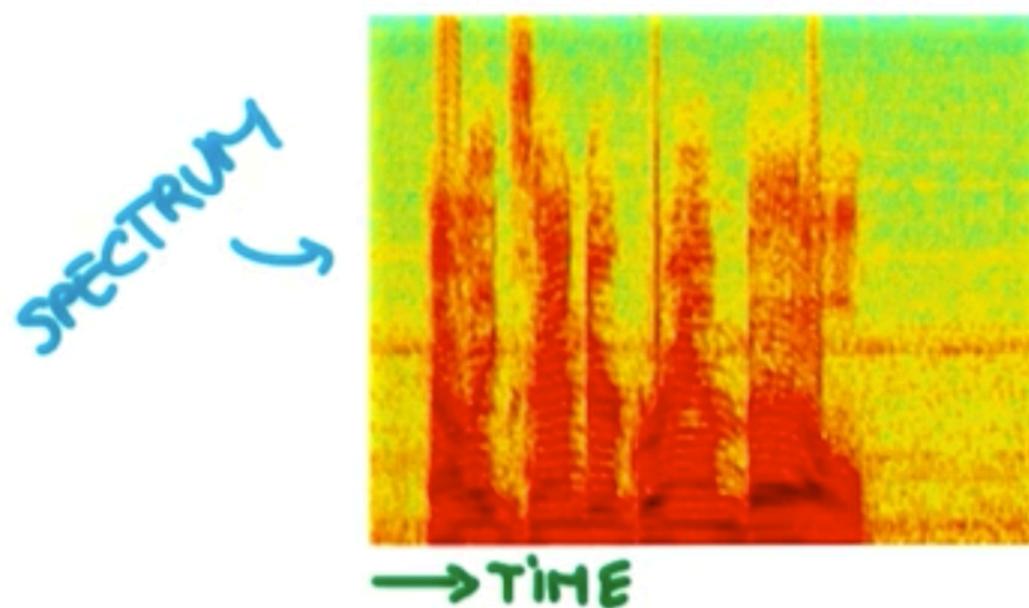
MACHINE TRANSLATION !

THE QUICK BROWN FOX → LE RENARD PRESSÉ

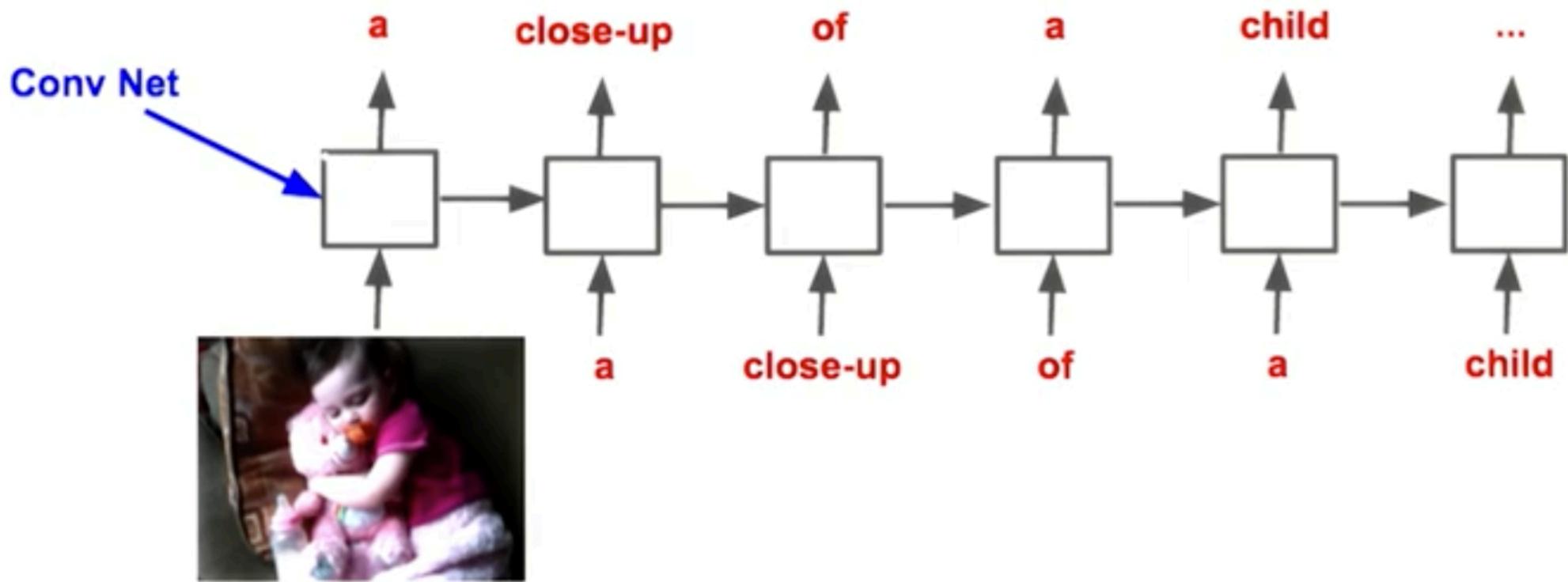


$$X_1, X_2, X_3 \rightarrow Y_1, Y_2, Y_3, Y_4$$

SPEECH RECOGNITION!



→ "HELLO WORLD"



Describes without errors

Describes with minor errors

Somewhat related to the image

Unrelated to the image



A person riding a motorcycle on a dirt road.



Two dogs play in the grass.



A skateboarder does a trick on a ramp.



A dog is jumping to catch a frisbee.



A group of young people playing a game of frisbee.



Two hockey players are fighting over the puck.



A little girl in a pink hat is blowing bubbles.



A refrigerator filled with lots of food and drinks.