Multi-Layer Neural Networks

NLP in one day

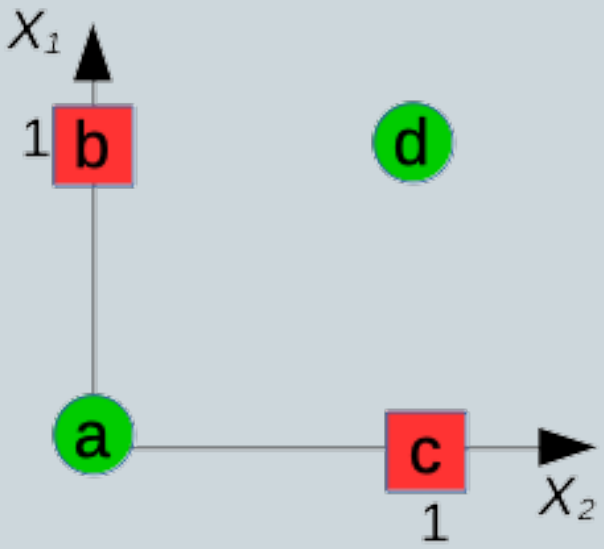
KING'S College LONDON

NIHR | Maudsley Biomedical Research Centre

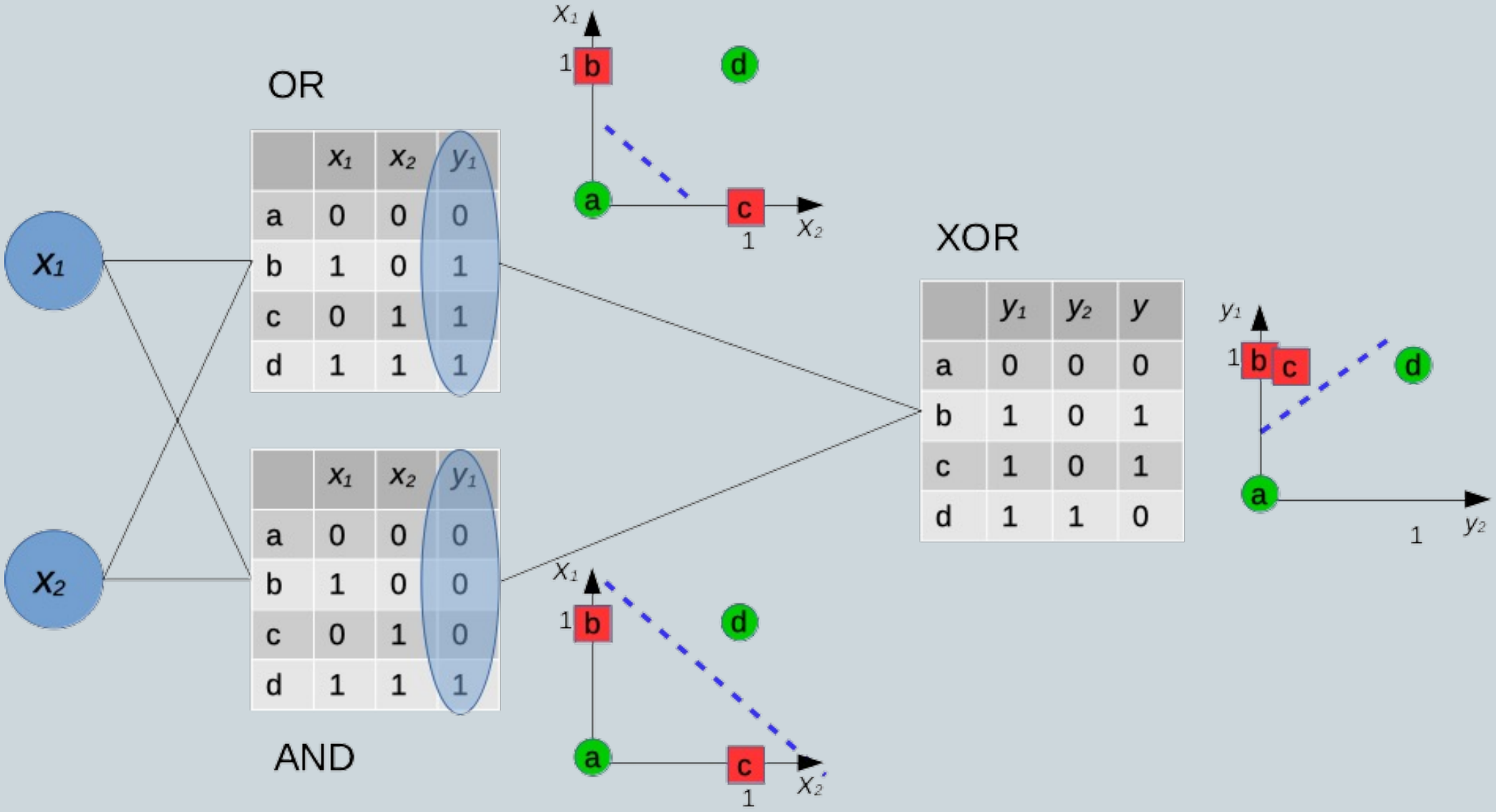
XOR is not linearly separable

XOR

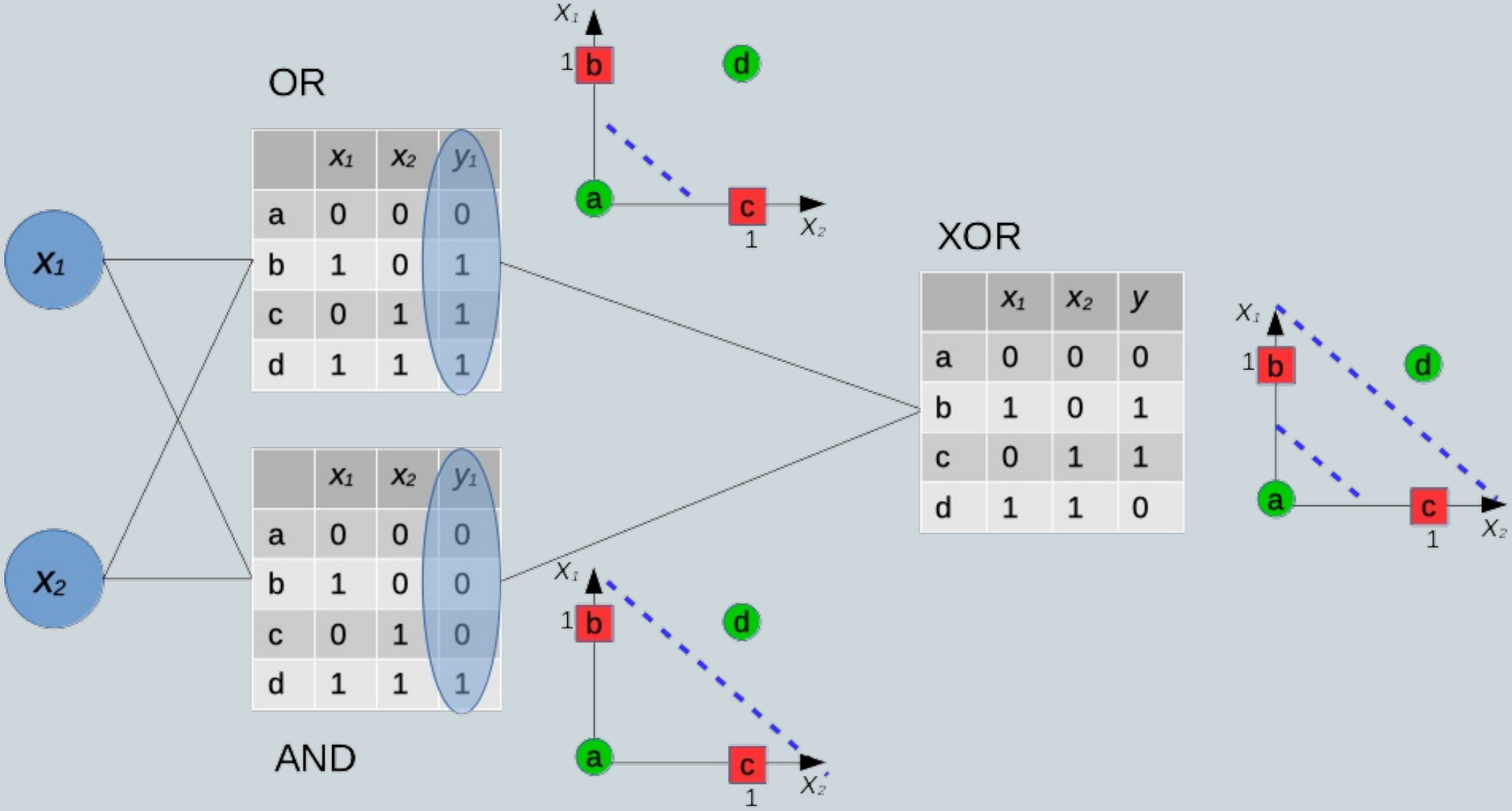
	x_1	x_2	y
a	0	0	0
b	1	0	1
c	0	1	1
d	1	1	0



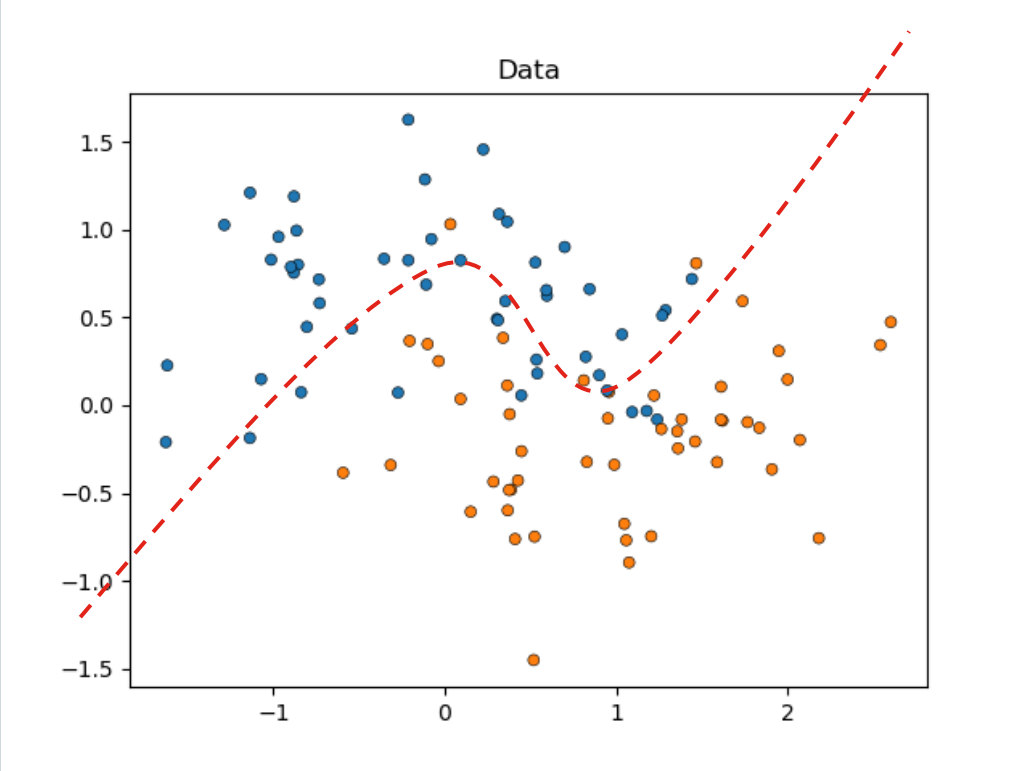
Solving a non-linear problem

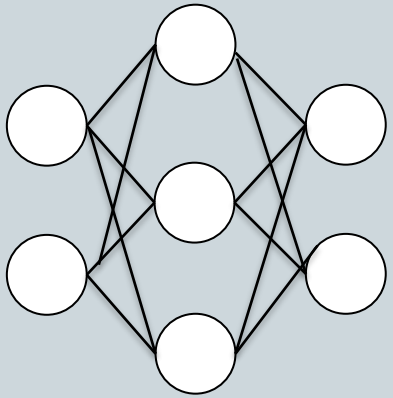


Solving a non-linear problem



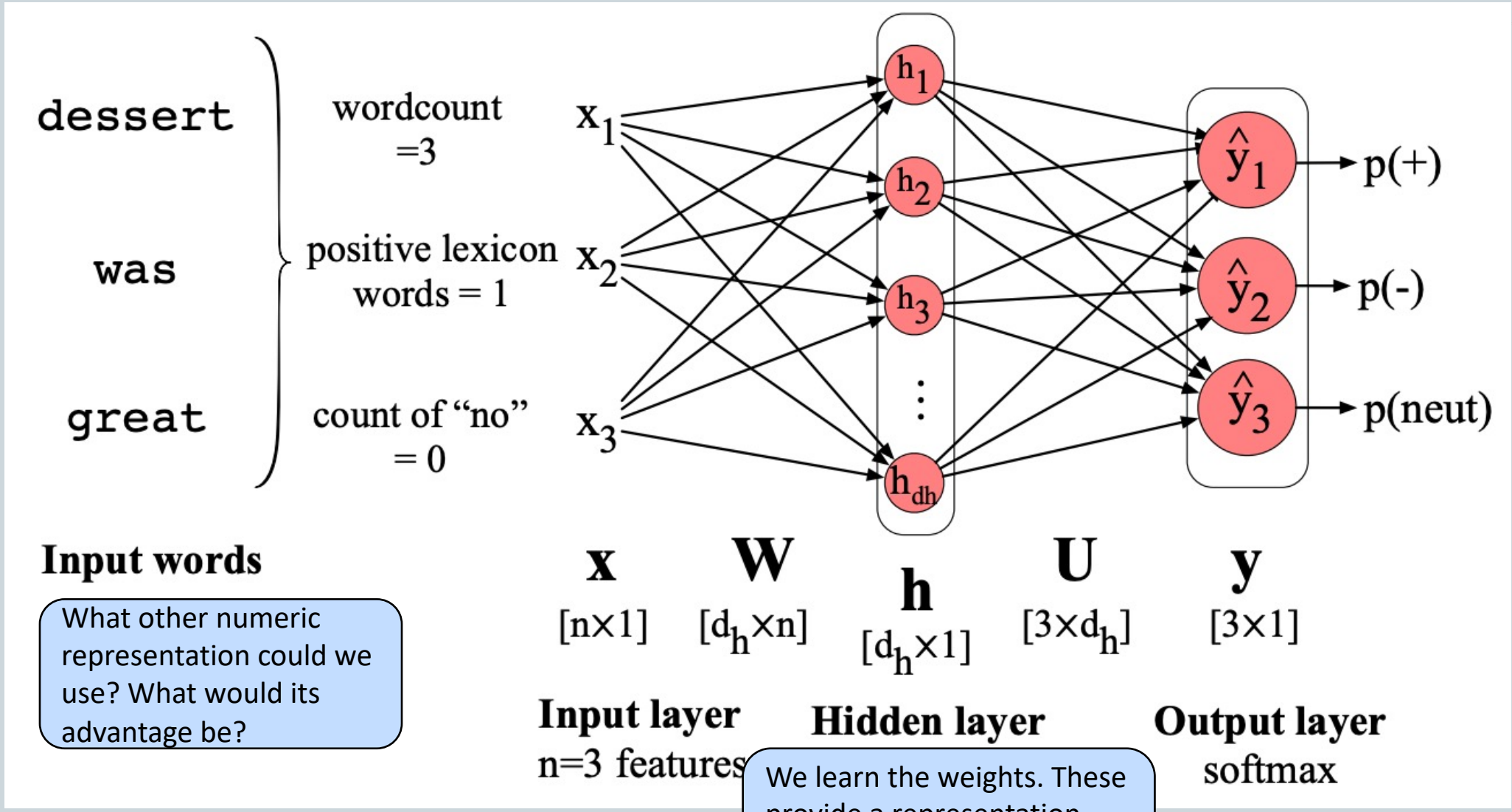
A single perceptron can only model linearly separable problems





Solvable with 12 parameters (weights)

How do we model language in a neural network?



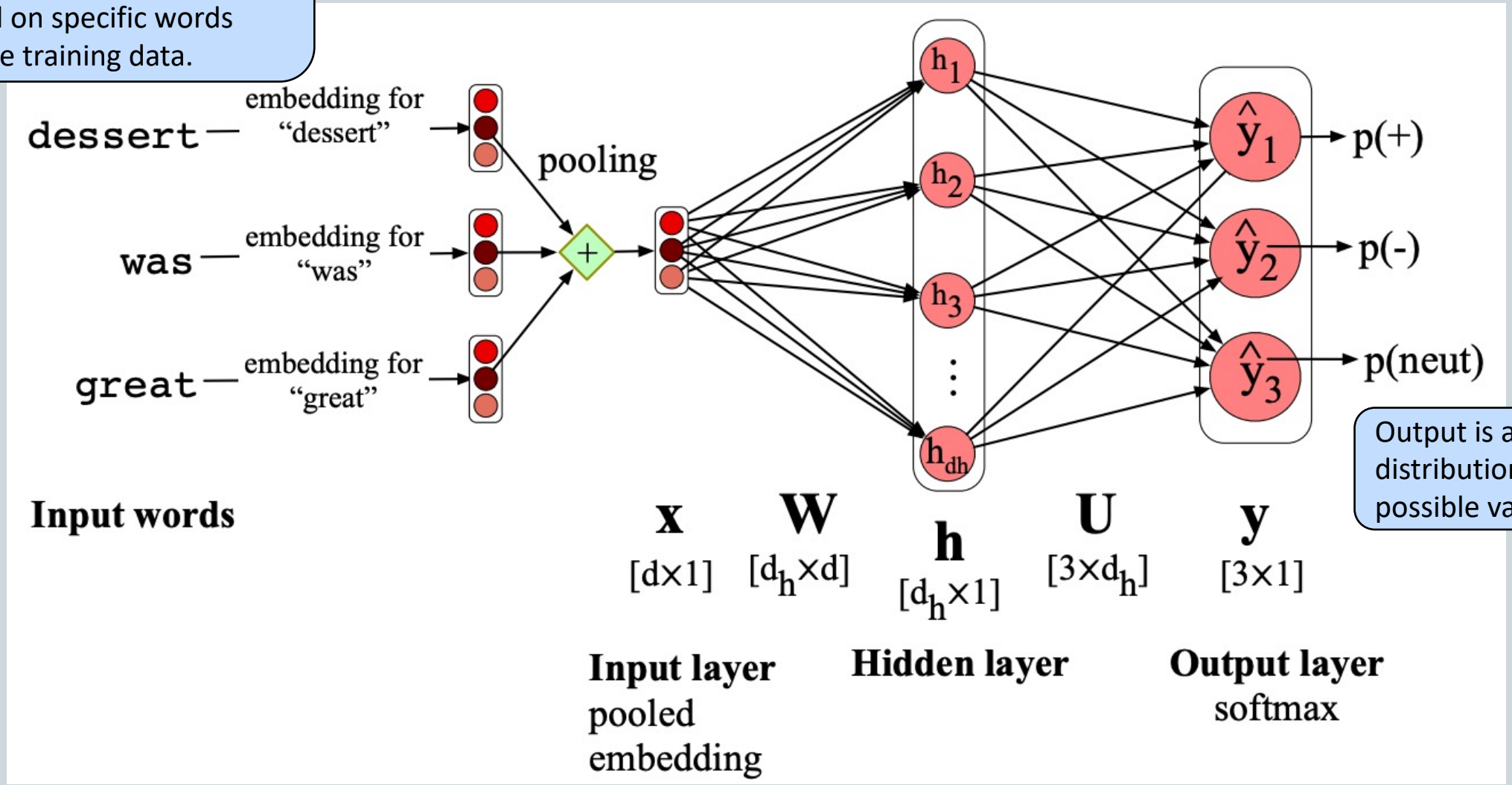
What other numeric representation could we use? What would its advantage be?

We learn the weights. These provide a representation that will map input to output.

(Jurafsky and Martin, Fig. 7.10)

How do we model language in a neural network?

Replace the words with embeddings. The model will be more general, rather than focused on specific words from the training data.



Output is a probability distribution over all possible values

(Jurafsky and Martin, Fig. 7.11)

Thank you

angus.roberts@kcl.ac.uk
https://www.kcl.ac.uk/people/angus-roberts