



Department of Computer Science

UNIVERSITY OF COLORADO **BOULDER**



Deep Learning

Advanced Machine Learning for NLP

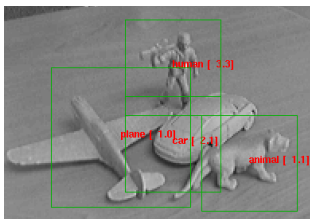
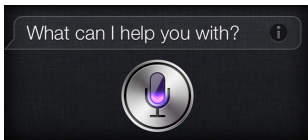
Jordan Boyd-Graber

INTRODUCTION

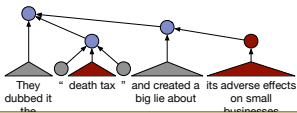
Deep Learning was once known as “Neural Networks”



But it came back ...



- More data
- Better tricks (regularization)
- Faster computers



And companies are investing ...

Google Hires Brains that Helped Supercharge Machine Learning

BY ROBERT MCMILLAN 03.13.13 | 6:30 AM | PERMALINK

[f Share](#) [0](#) [t Tweet](#) [1](#) [g+1](#) [145](#) [in Share](#) [Pintt](#)



And companies are investing ...

'Chinese Google' Opens Artificial-Intelligence Lab in Silicon Valley

BY DANIELA HERNANDEZ 04.12.13 | 6:30 AM | [PERMALINK](#)

[f Share](#) 0 [t Tweet](#) 0 [g+1](#) 228 [in Share](#) [Pintt](#)



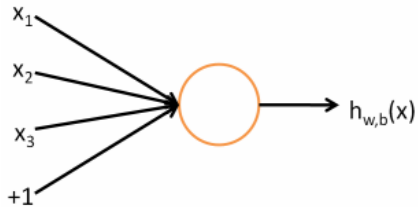
And companies are investing ...

Facebook's 'Deep Learning' Guru Reveals the Future of AI

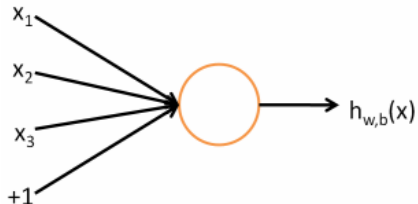
BY CADE METZ 12.12.13 | 6:30 AM | PERMALINK



Map inputs to output



Map inputs to output

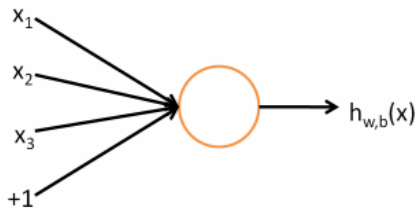


Input

Vector $x_1 \dots x_d$

inputs encoded as
real numbers

Map inputs to output



Input

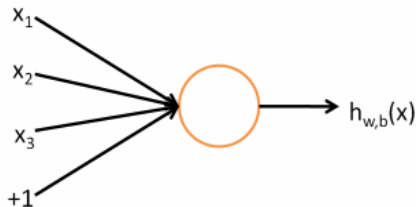
Vector $x_1 \dots x_d$

Output

$$f\left(\sum_i W_i x_i + b\right)$$

multiply inputs by weights

Map inputs to output



Input

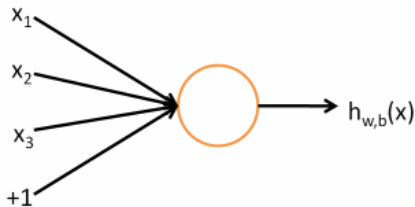
Vector $x_1 \dots x_d$

Output

$$f\left(\sum_i W_i x_i + b\right)$$

add bias

Map inputs to output



Input

Vector $x_1 \dots x_d$

Output

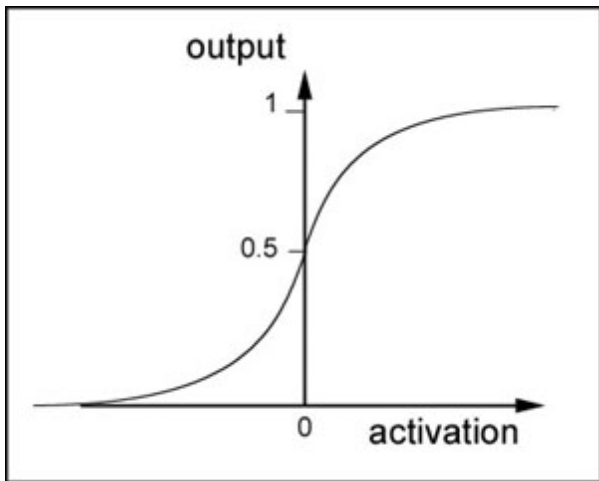
$$f\left(\sum_i W_i x_i + b\right)$$

Activation

$$f(z) \equiv \frac{1}{1 + \exp(-z)}$$

pass through
nonlinear sigmoid

Why is it called activation?



In the shallow end

- This is still logistic regression
- Engineering features x is difficult (and requires expertise)
- Can we learn how to represent inputs into final decision?