1- introduction

deep learning = branch / subfield of machine learning that involves the use of neural networks to model and solve complex problems

↳ deep ⟹ implies more than a few layers

$$\hat{y} = f(x; \Theta)$$

output ↲   ↓ model   ↳ parameters   → input

$f(x_1) \rightarrow y_1$
$f(x_2) \rightarrow y_2$ $f(x_1 + x_2) \rightarrow (y_1 + y_2)$

- any function can be approximated using superpositions of a sigmoidal function

sigmoid function = $\sigma(x) = \dfrac{1}{1+e^{-x}}$ $\overbrace{\sigma(-x), \sigma(\frac{x}{0.01}), \sigma(x)+2}$

history = perceptron ⟶ backpropagation ⟶ SVM ⟶ DL

- stupid ways to initialize weights
- using wrong type of non-linearity   } why backpropagation failed back then
- small datasets
- slow computers