How Deep Learning Works

SEOULTECH 국립서울과학기술대학교

Sangheum Hwang

Learning representations

- Learning representations from data
  - It is about the difference between deep learning and other ML approaches.
  - What ML algorithms do: to do ML, we need three things;
    - Input data points
    - Examples of the expected output
    - A way to measure whether the algorithm is doing a good job → adjusting the measurement is what we call learning.
  - The central problem of ML: how to meaningfully transform data → how to learn useful *representations* of the input data

Learning representations

- What is a representation?
  - It is a different way to look at data – to represent or encode data.
    - E.g., a color image → RGB format or HSV format
  - We can solve some tasks very easily with a good representation.
    - E.g., “select all red pixels” → very easy in the RGB format
  - ML models are all about finding appropriate representations for their input data!
    - Searching for useful representations of some input data, within a predefined space of possibilities, using guidance from a feedback signal.

Learning representations

- What is a representation?
  - Example
    - We want to develop an algorithm that can take (x,y) of a point and output whether that point is likely to be black or to be white.
    - Inputs? (x,y)
    - The expected outputs? 1 or 0 Black or White
    - How to measure the performance? accuracy

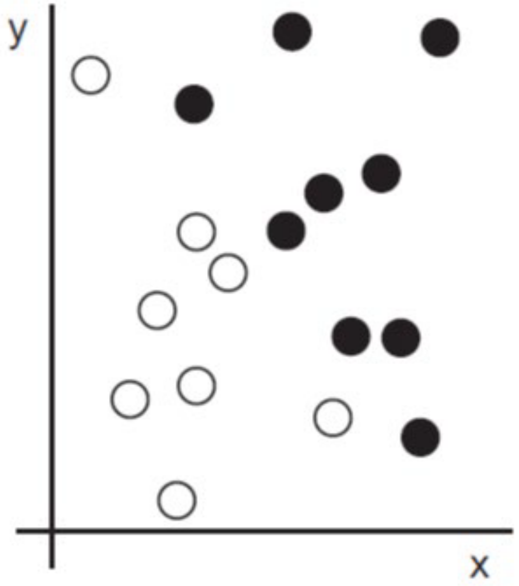


Figure 1.3
Some sample data

Learning representations

- What is a representation?
  - Example

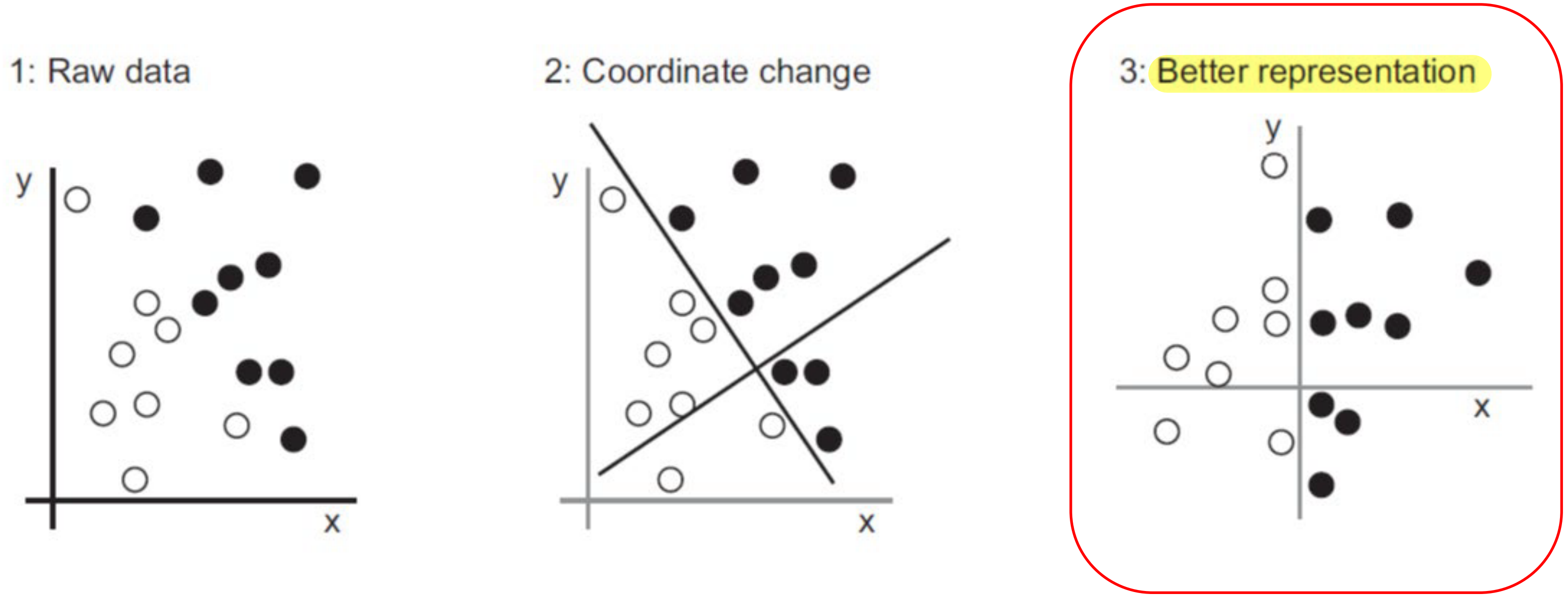


Figure 1.4 Coordinate change

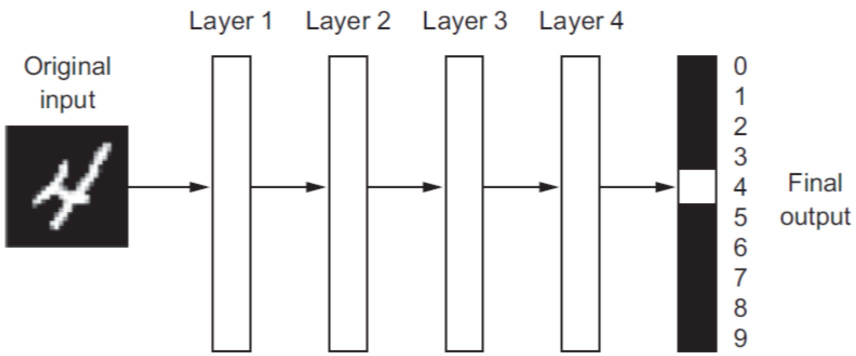
If x>0, then a point is black
If x<0, then a point is white

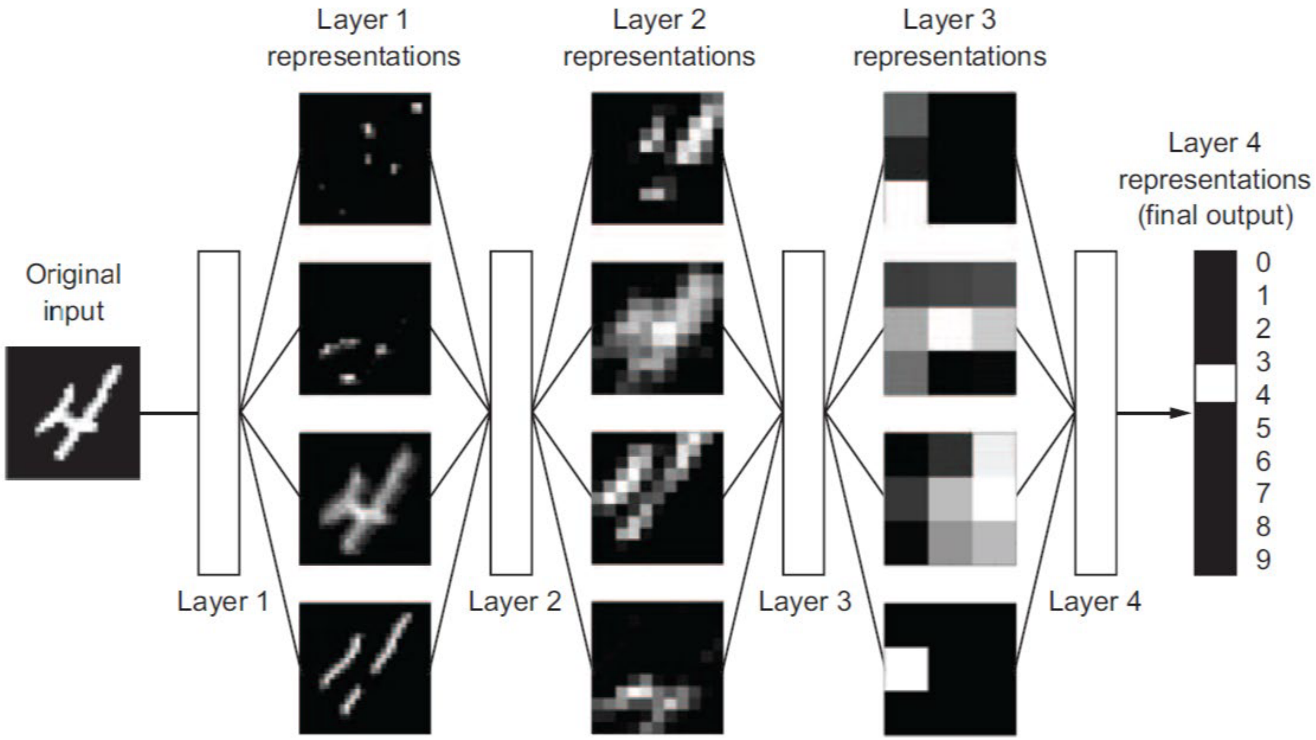
The “deep” in “deep learning”

- Deep learning
  - A new take on learning representations from data by learning successive layers.
  - “deep” stands for this idea of successive layers of representations.
  - Modern deep learning often involves tens of even hundreds of successive layers of representations – all learned automatically from training data.
  - In DL, these layered representations are learned via neural networks model.
  - Deep learning models are not models of the brain.

The “deep” in “deep learning”

- Learned representations by a deep learning model

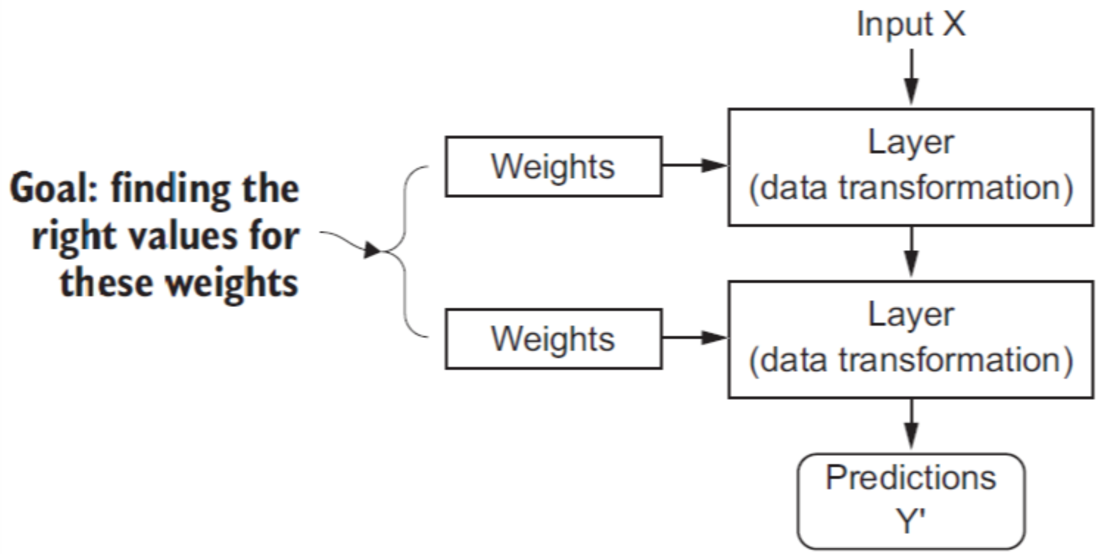




a multistage *information-distillation* process

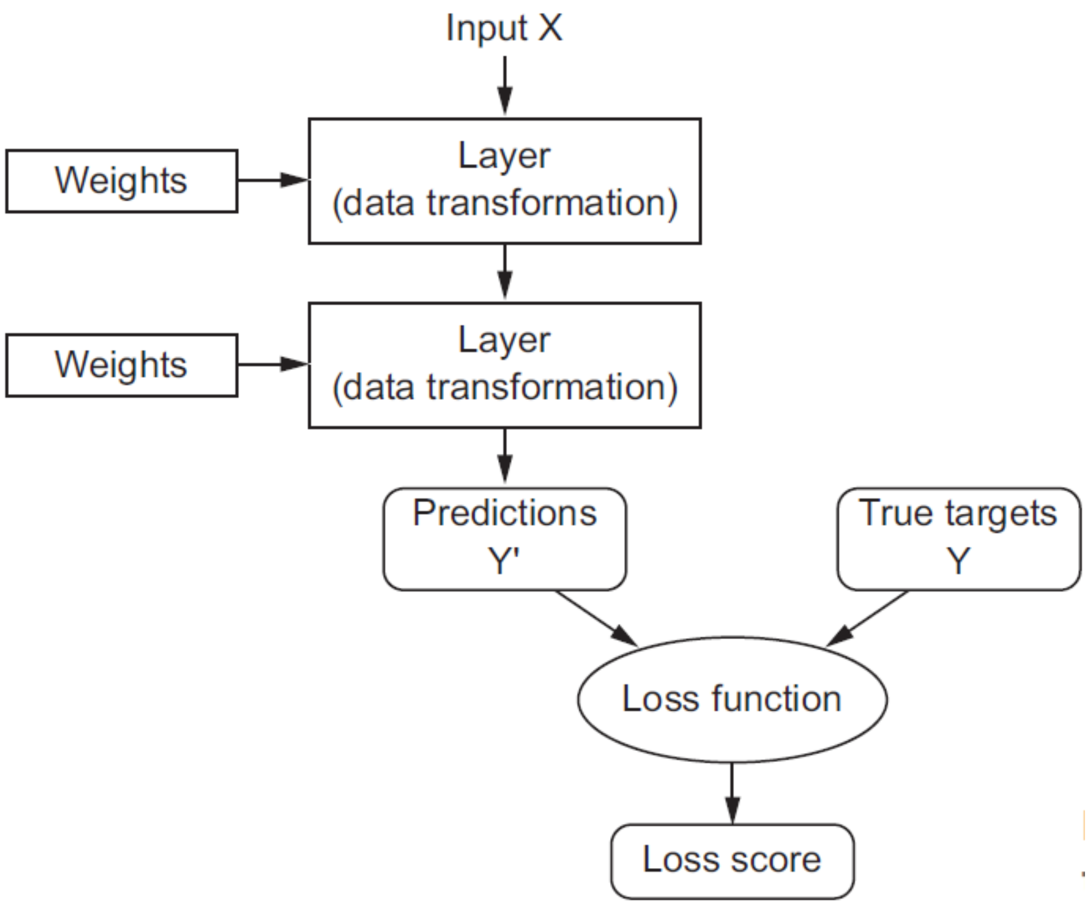
How deep learning works

- What a layer does to its input data is stored in the layer's *weights*.
- In other words, the transformation (by a layer) is *parameterized* by its weights.
- Again, learning means finding a set of weight values of all layers in a network.



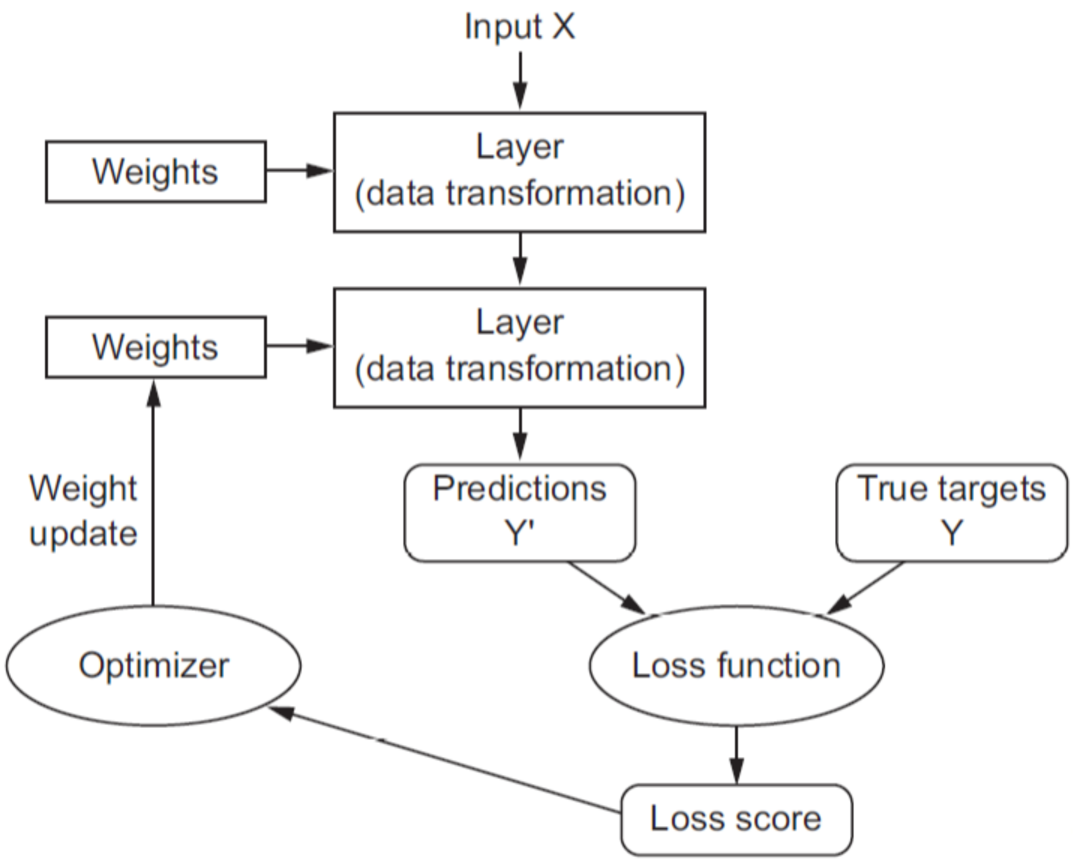
How deep learning works

- To control something, we need to observe it!
- To control the output of a neural network, we need to measure how far this output is from what we expected. → the *loss function* of the network (also called the objective function)
- Loss(output, target) → how well the network performs our task



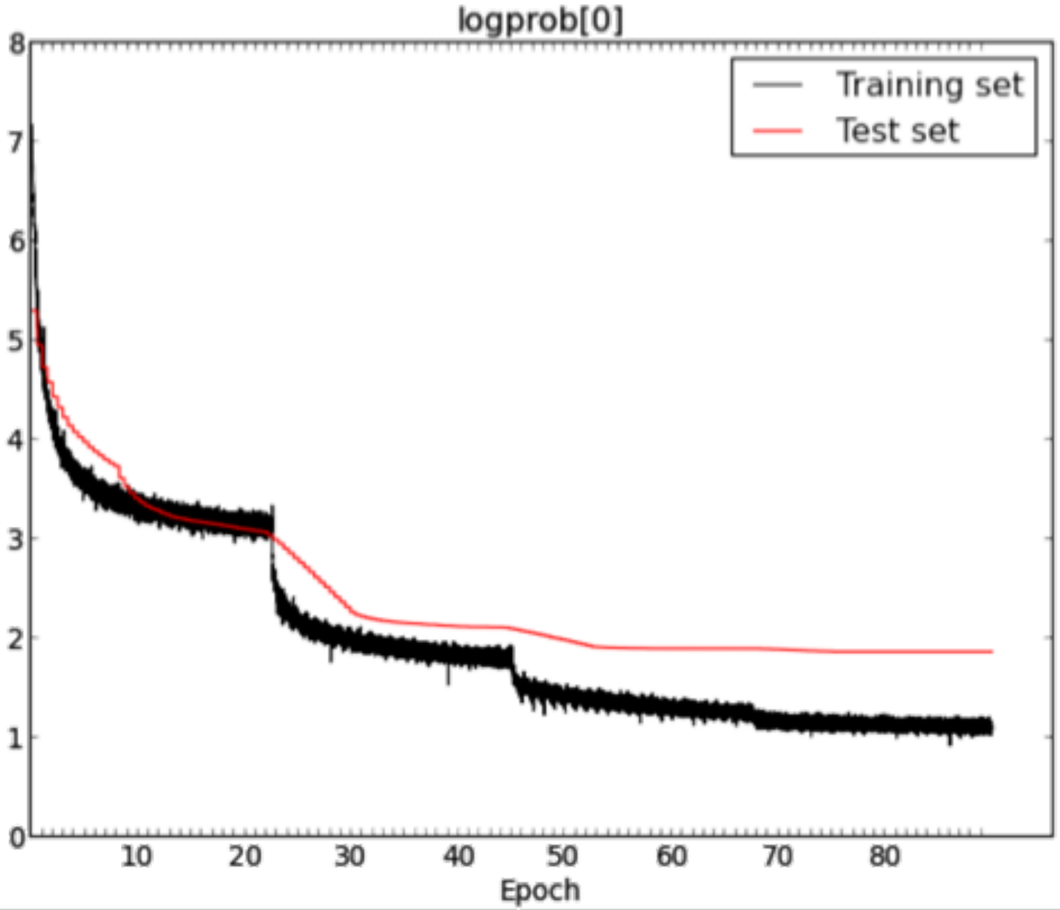
How deep learning works

- Deep learning adjusts the weight values a little, in a direction that will lower the loss score for the current example. → *Optimizer*'s job
- Specifically, the optimizer is based on the *backpropagation* algorithm.



How deep learning works

- Loss curves



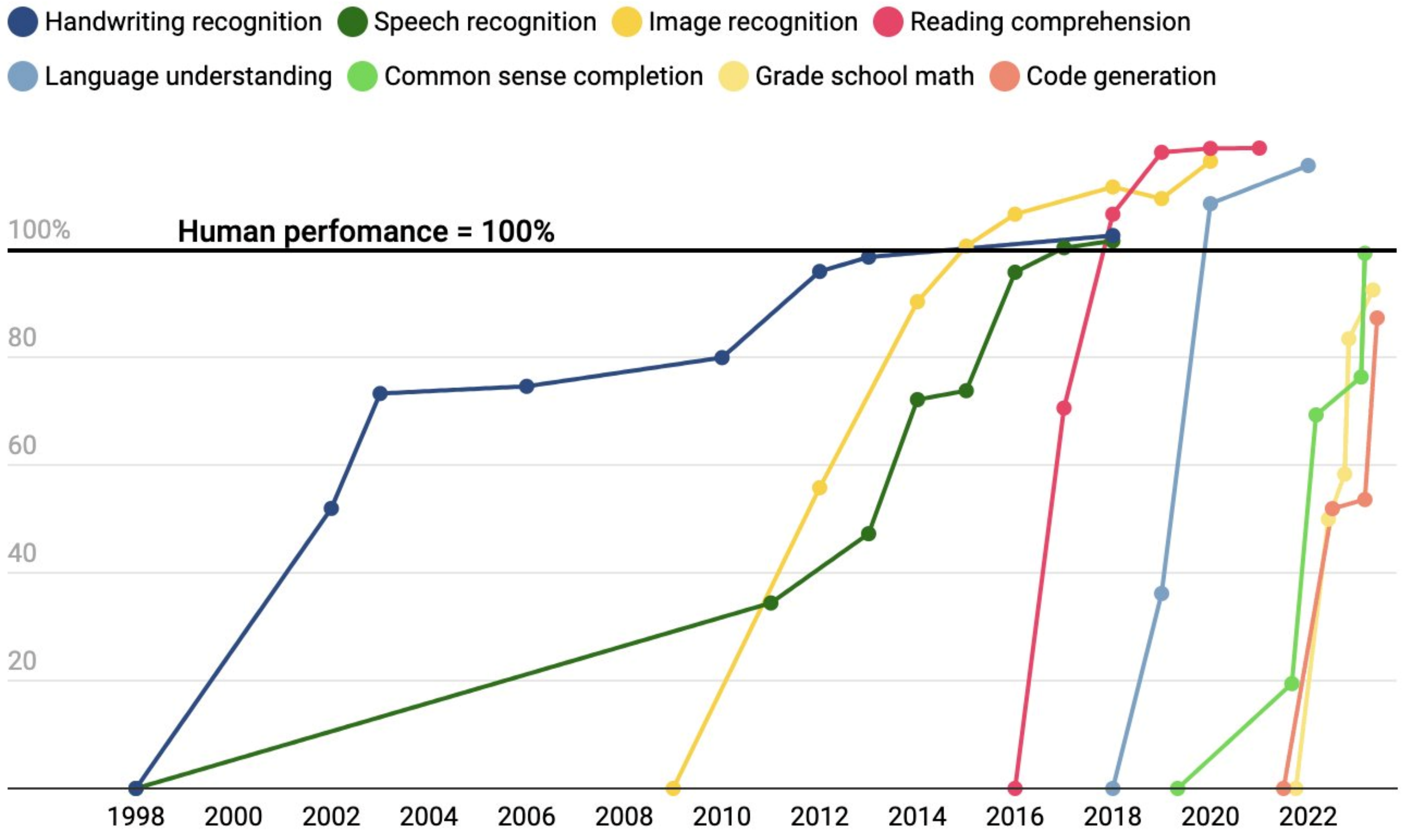
How deep learning works

- Learning process
  1) Initially, the weights of the network are assigned random values.
  2) With every example the network processes, the weights are adjusted a little in the correct direction. → The loss scores decreases
  3) Repeat a sufficient number of times, then we can get weight values that minimize the loss function.
  4) This network with a minimal will produce the outputs which are close to the targets. → Trained network

What DL has achieved

AI has surpassed humans at a number of tasks and the rate at which humans are being surpassed at new tasks is increasing

State-of-the-art AI performance on benchmarks, relative to human performance



For each benchmark, the maximally performing baseline reported in the benchmark paper is taken as the "starting point", which is set at 0%. Human performance number is set at 100%. Handwriting recognition = MNIST, Language understanding = GLUE, Image recognition = ImageNet, Reading comprehension = SQuAD 1.1, Reading comprehension = SQuAD 2.0, Speech recognition = Switchboard, Grade school math = GSK8k, Common sense completion = HellaSwag, Code generation = HumanEval.

Chart: Will Henshall for TIME • Source: ContextualAI

TIME