Advantages of classical ML

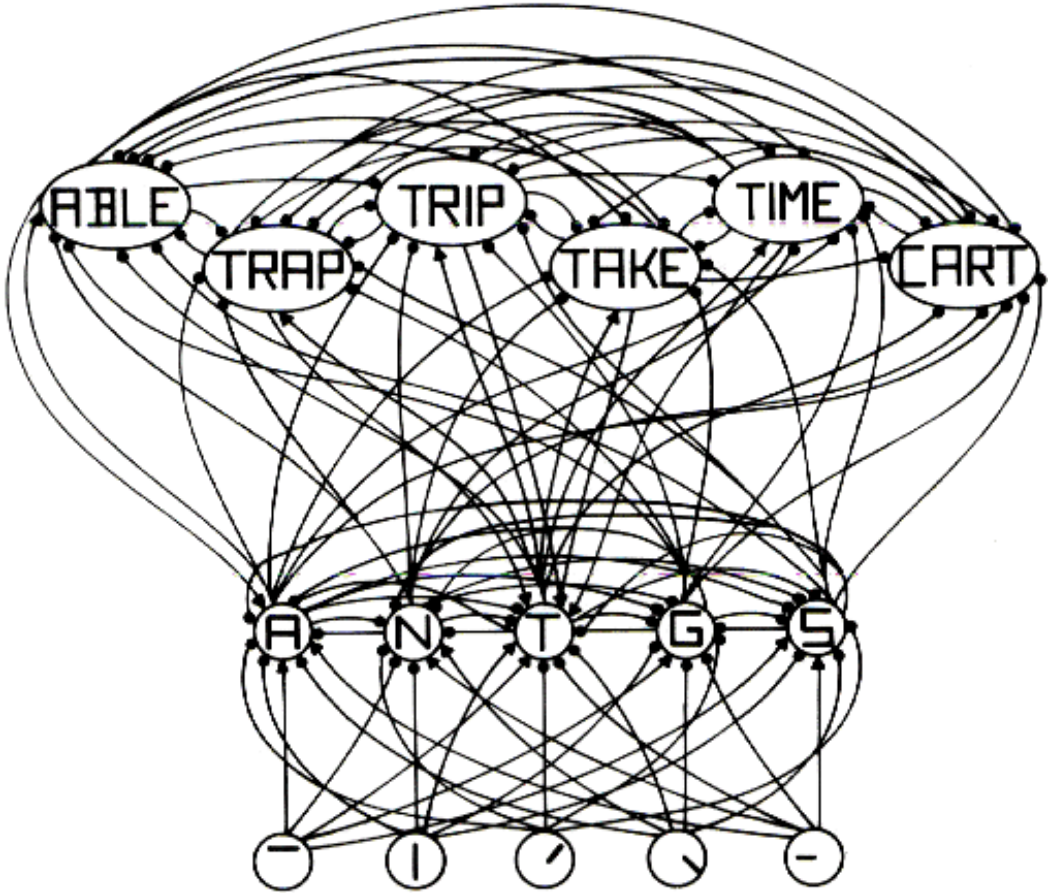
- Potential learning from few examples
- Emphasis on analytical (mathematical) results, e.g.

  - How many examples does a given target concept require to learn?

  - How expensive (time, space) is the computation (computational complexity)?

- Disadvantages: Doesn't work (yet)

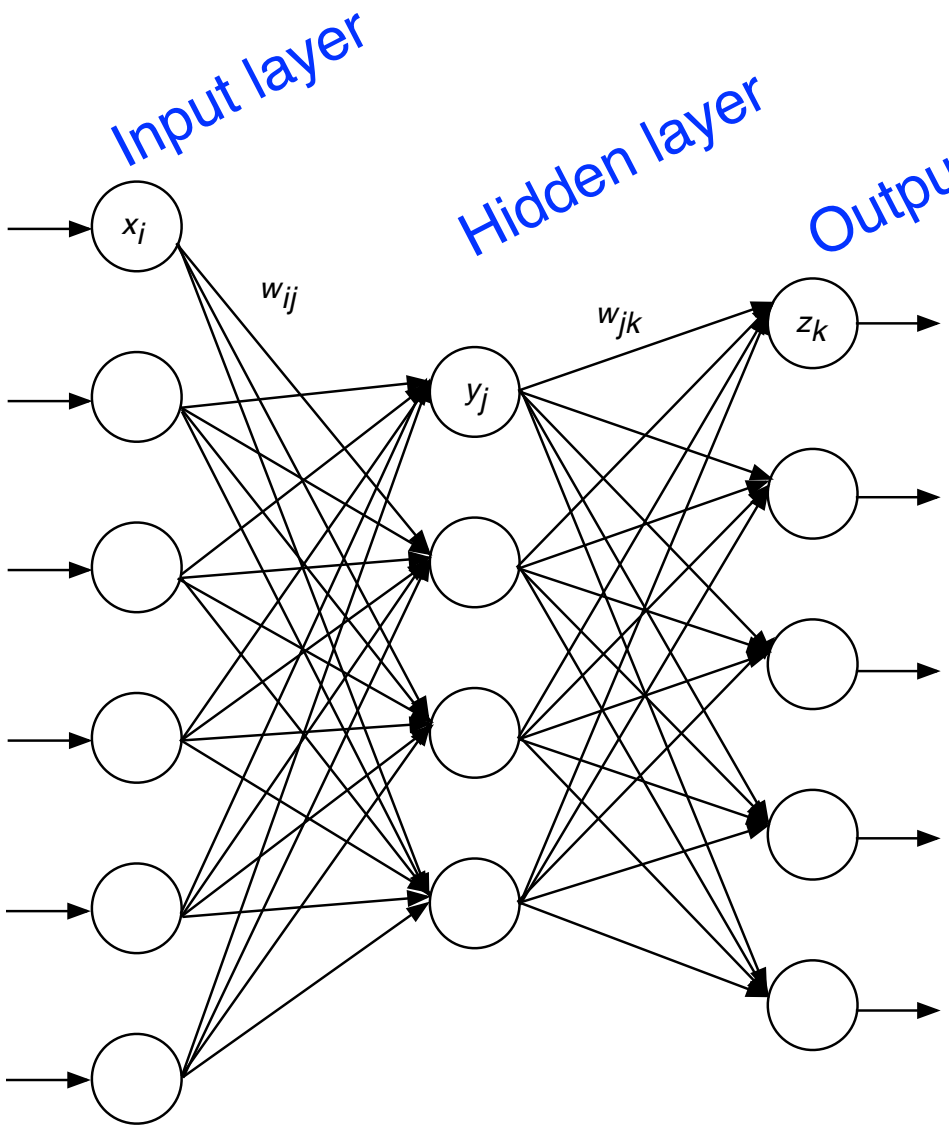
McClelland & Rumelhart's "Interactive Activation" model of reading (1981)

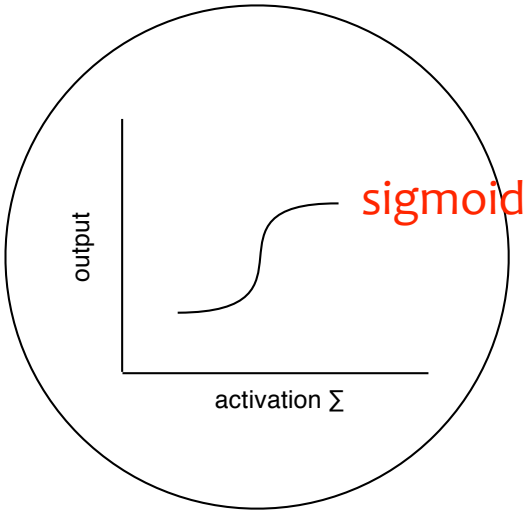


Output

Input

Multi-layer perceptrons





Connectionism in the 1980s
aka Parallel Distributed Processing (1986)

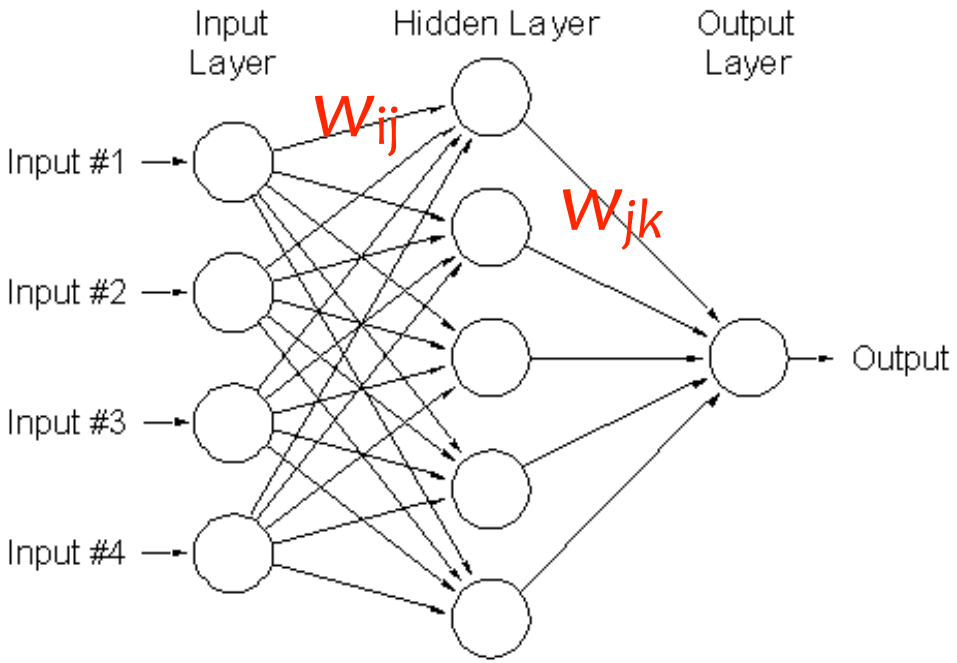
Parallel: Many nodes that all run at the same time.
Distributed: Knowledge is represented as the weights on the connections

Backpropagation algorithm:

1. Feed an input through the network; obtain an output.

2. An "oracle" compares the actual output to the target output (*supervised learning*)

3. Using the "error" (discrepancy), update the weights on all the connections.



w_{ij} = weight on connection from node i to node j

PDP/connectionist learning

- The model is biologically plausible or at least biologically inspired
- There is one general principle of learning: update the weights in the network in light of experience

  - Modern analog of empiricist "formation of associations"

- All knowledge is distributed among the connections

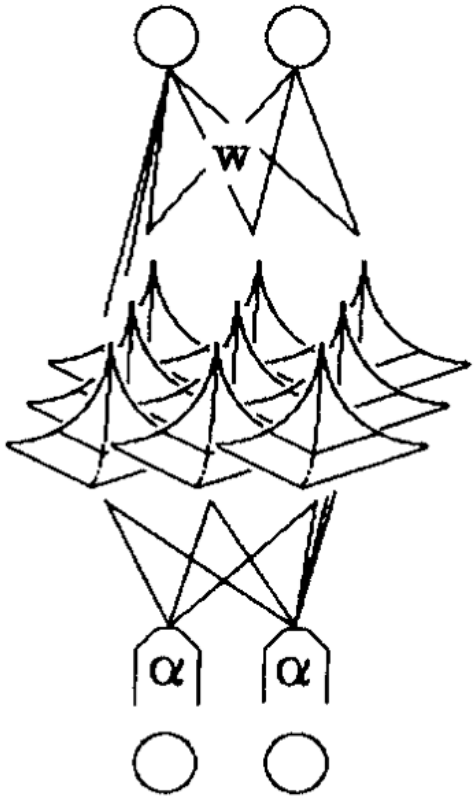
  - There are no symbols, separate concepts, etc.

Exemplar-based categorization in a neural network

Same basic idea applied to categorization!

Nodes represent exemplars.

Changing the weights (learning) means changing which stimulus features tend to active which exemplars and thus which categories.



Category nodes.

Learned association weights.

Exemplar nodes.

Learned attention strengths.

Stimulus dimension nodes.

The ALCOVE model of categorization(1996)