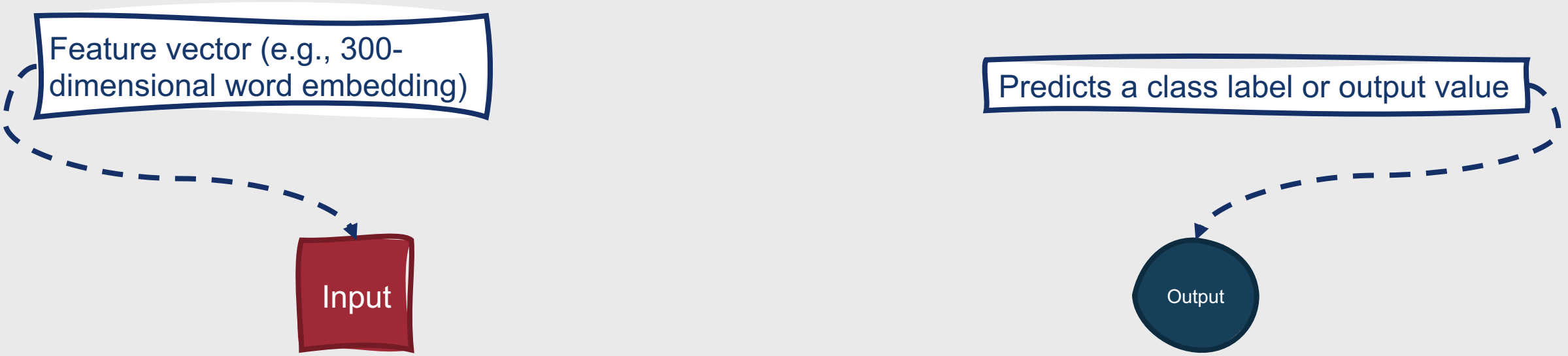
Feedforward Neural Network Basics

Natalie Parde
UIC CS 421

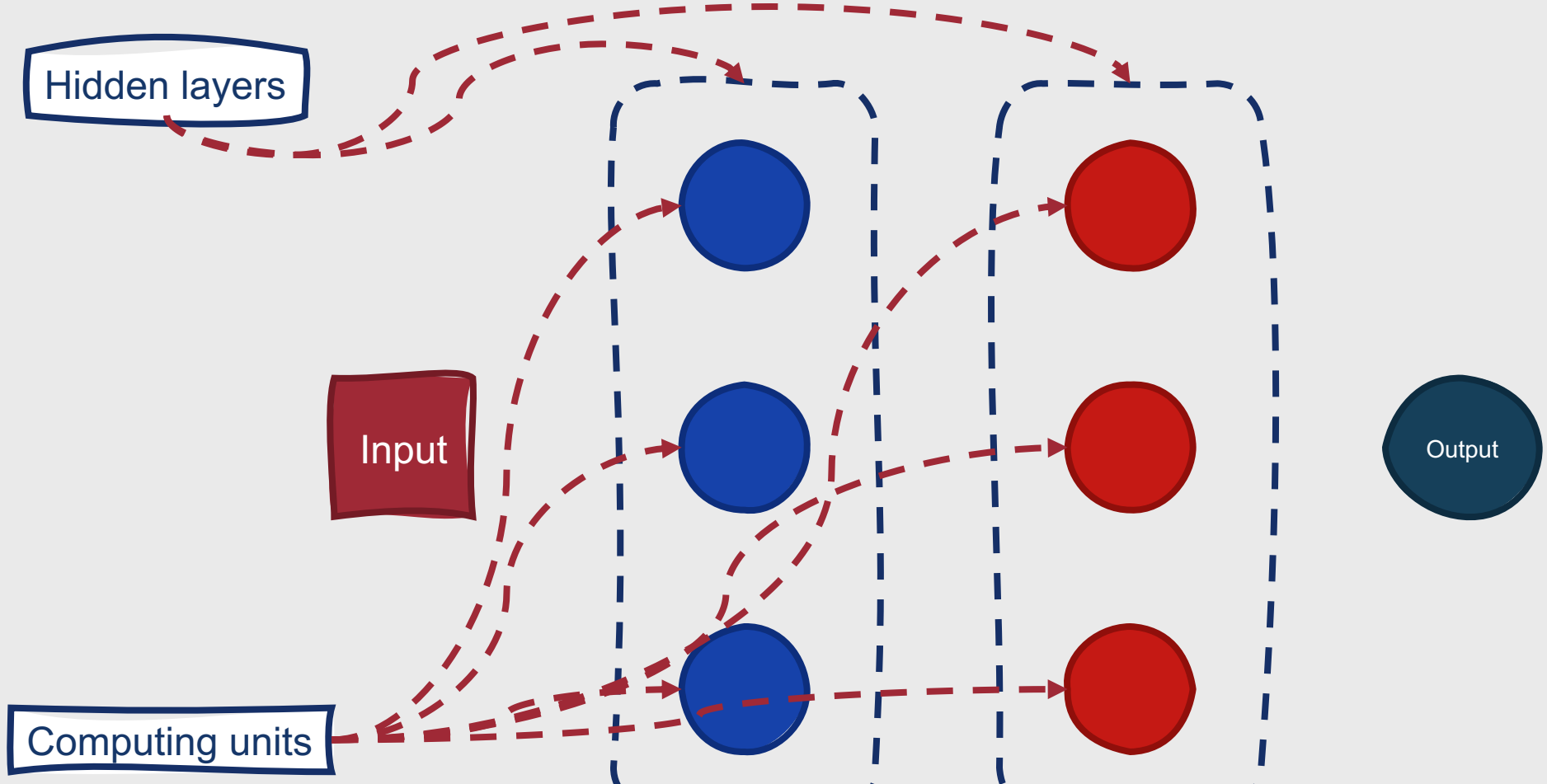
Feedforward Neural Networks

- Earliest and simplest form of neural network
- Data is fed forward from one layer to the next
- Each layer:
  - One or more units
  - A unit in layer *n* receives input from all units in layer *n*-1 and sends output to all units in layer *n*+1
  - A unit in layer *n* does not communicate with any other units in layer *n*
- The outputs of all units except for those in the last layer are **hidden** from external viewers

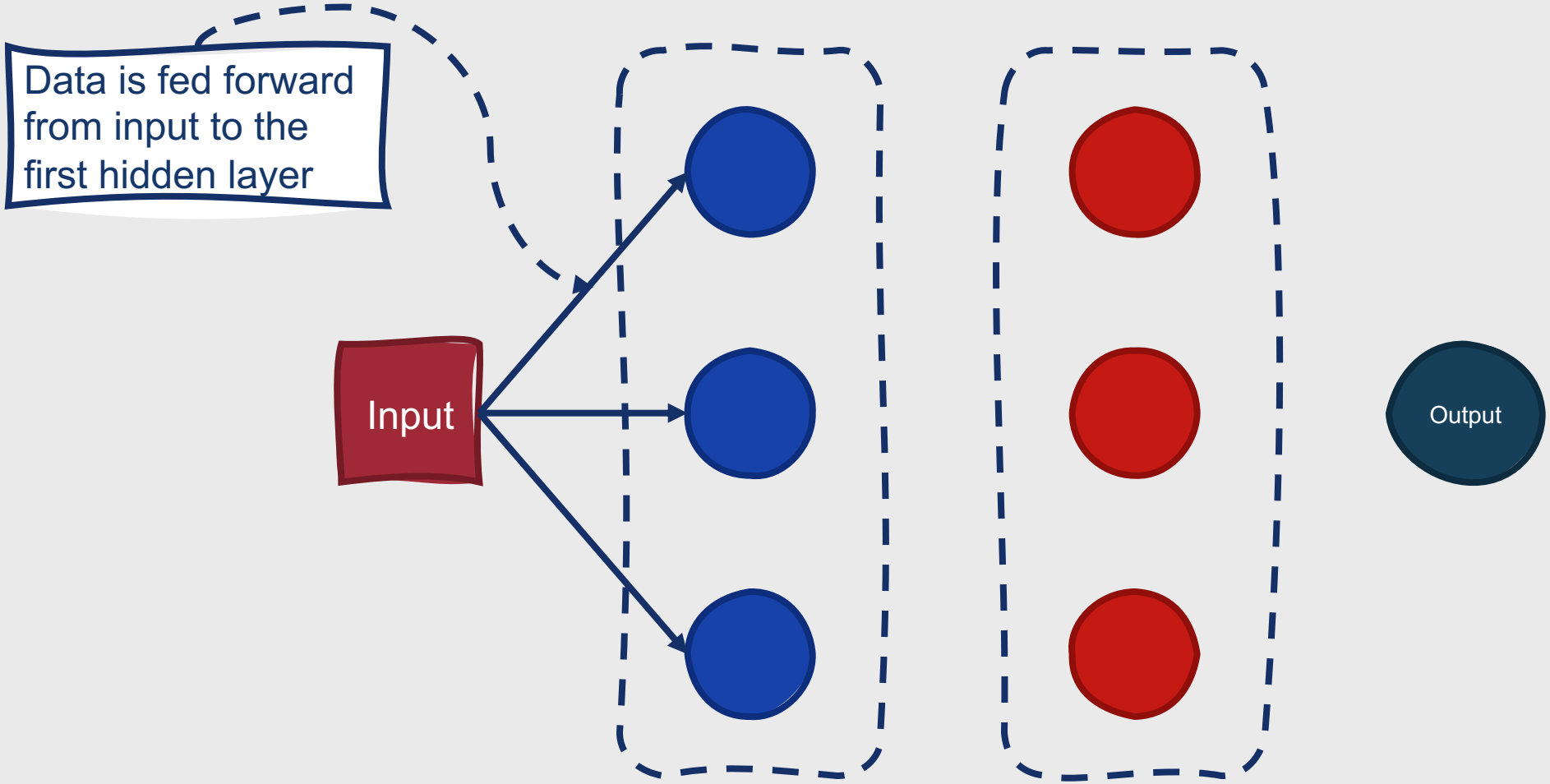
Feedforward Neural Networks



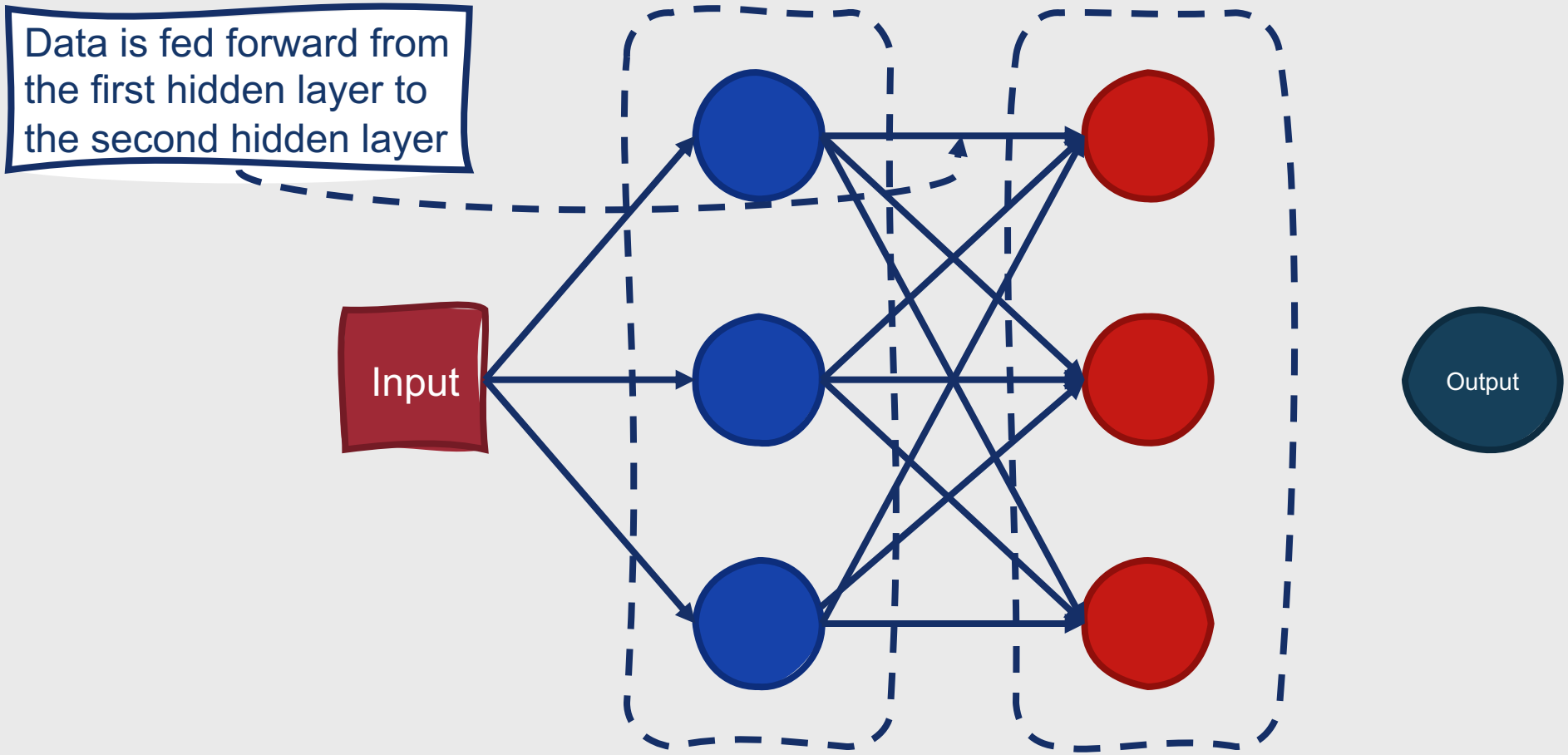
Feedforward Neural Networks



Feedforward Neural Networks

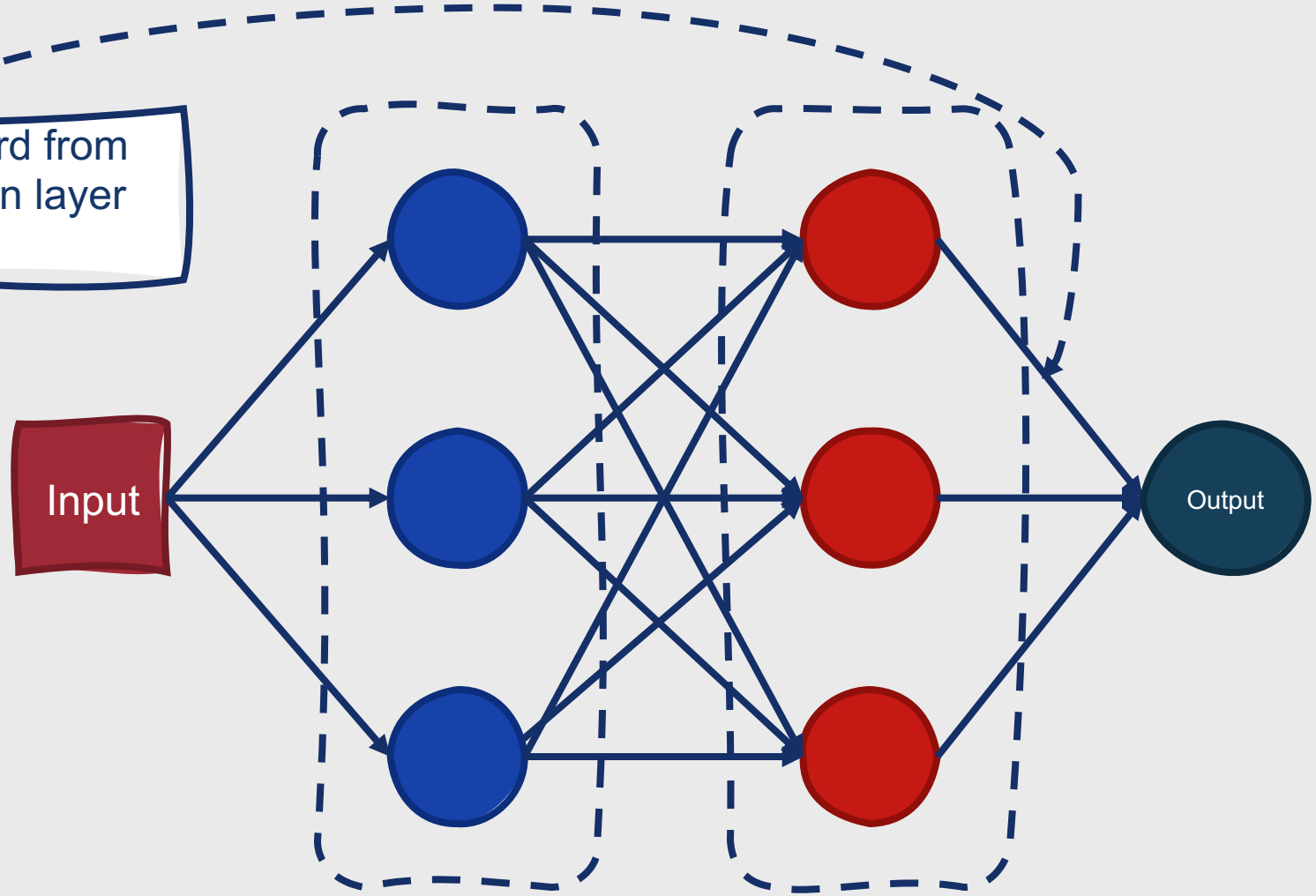


Feedforward Neural Networks

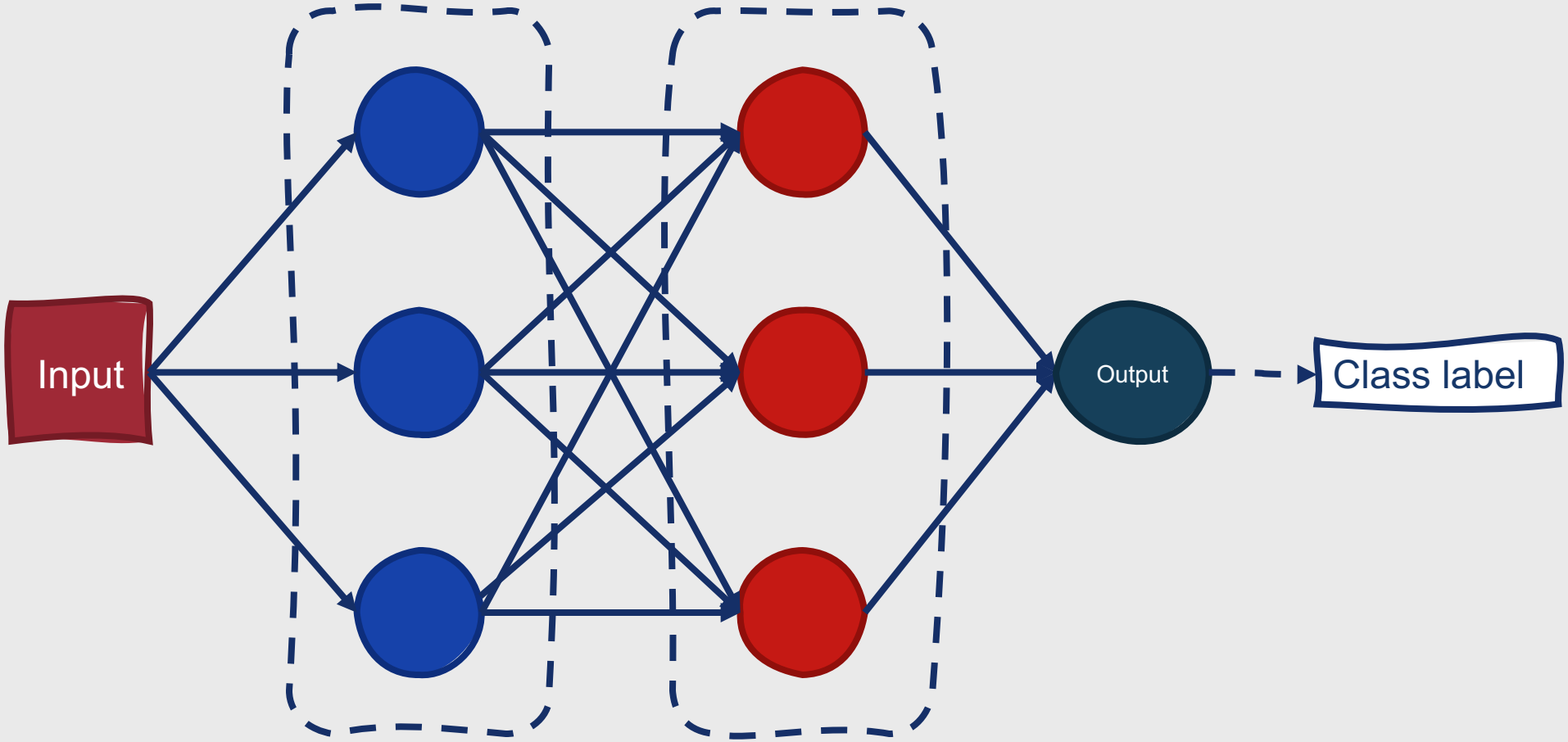


Feedforward Neural Networks

Data is fed forward from the second hidden layer to the output unit



Feedforward Neural Networks



Are feedforward neural networks an example of deep learning?

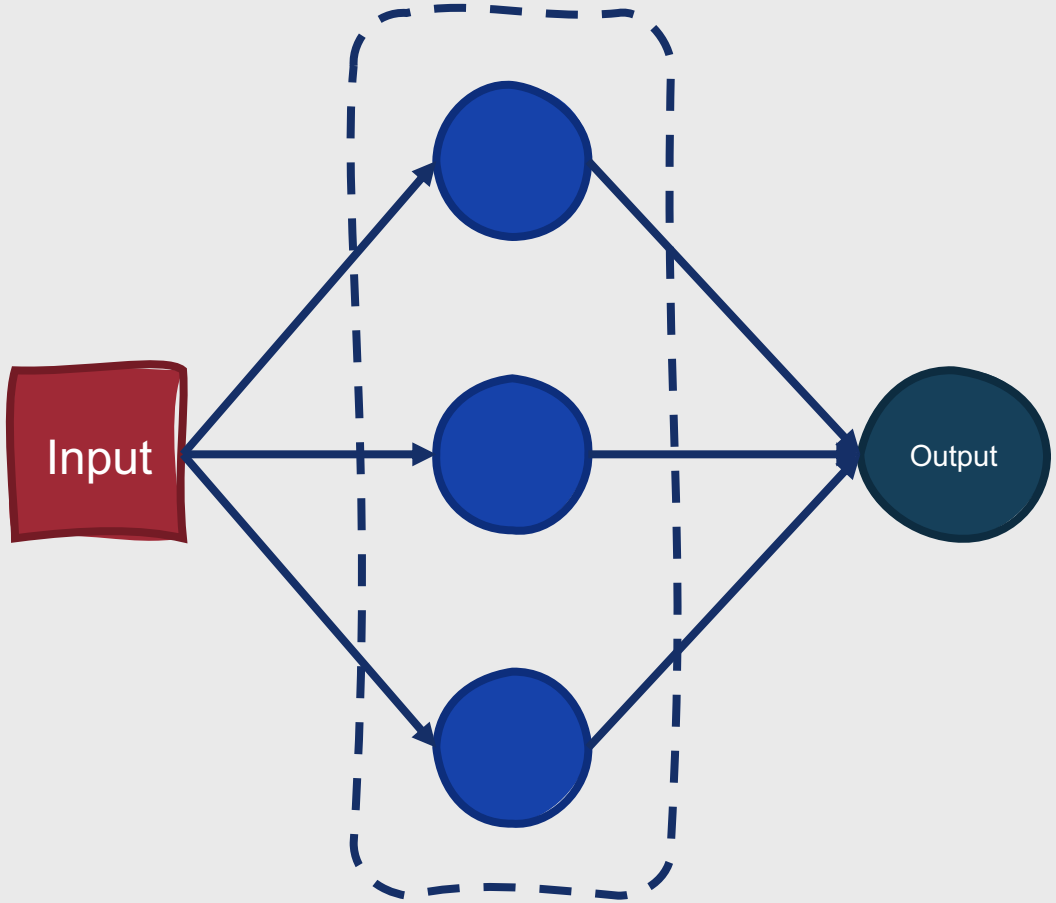
Yes ...if they have multiple layers

People often tend to refer to neural network-based machine learning as **deep learning**

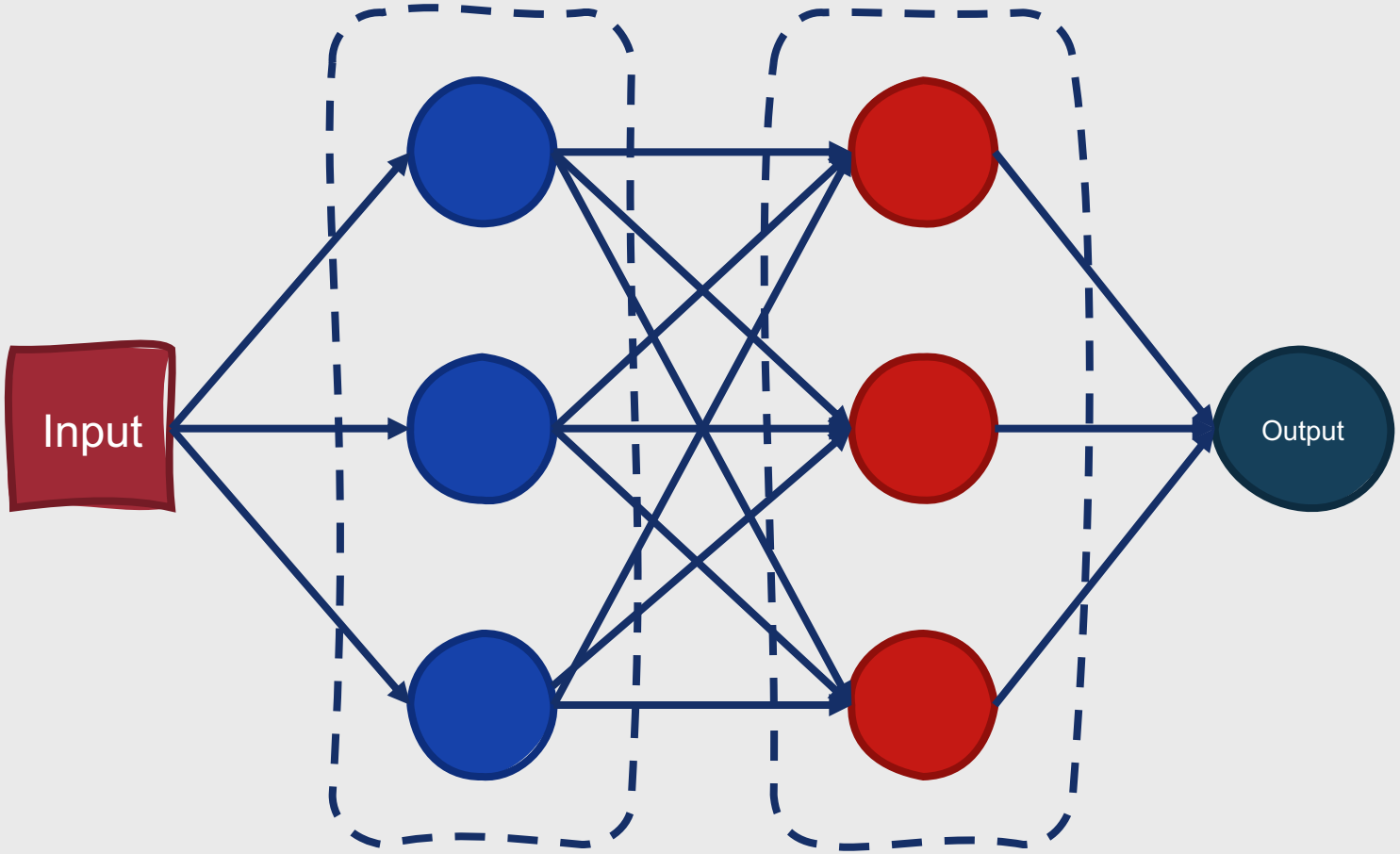
Why?

- Modern networks often have many layers (in other words, they're **deep**)

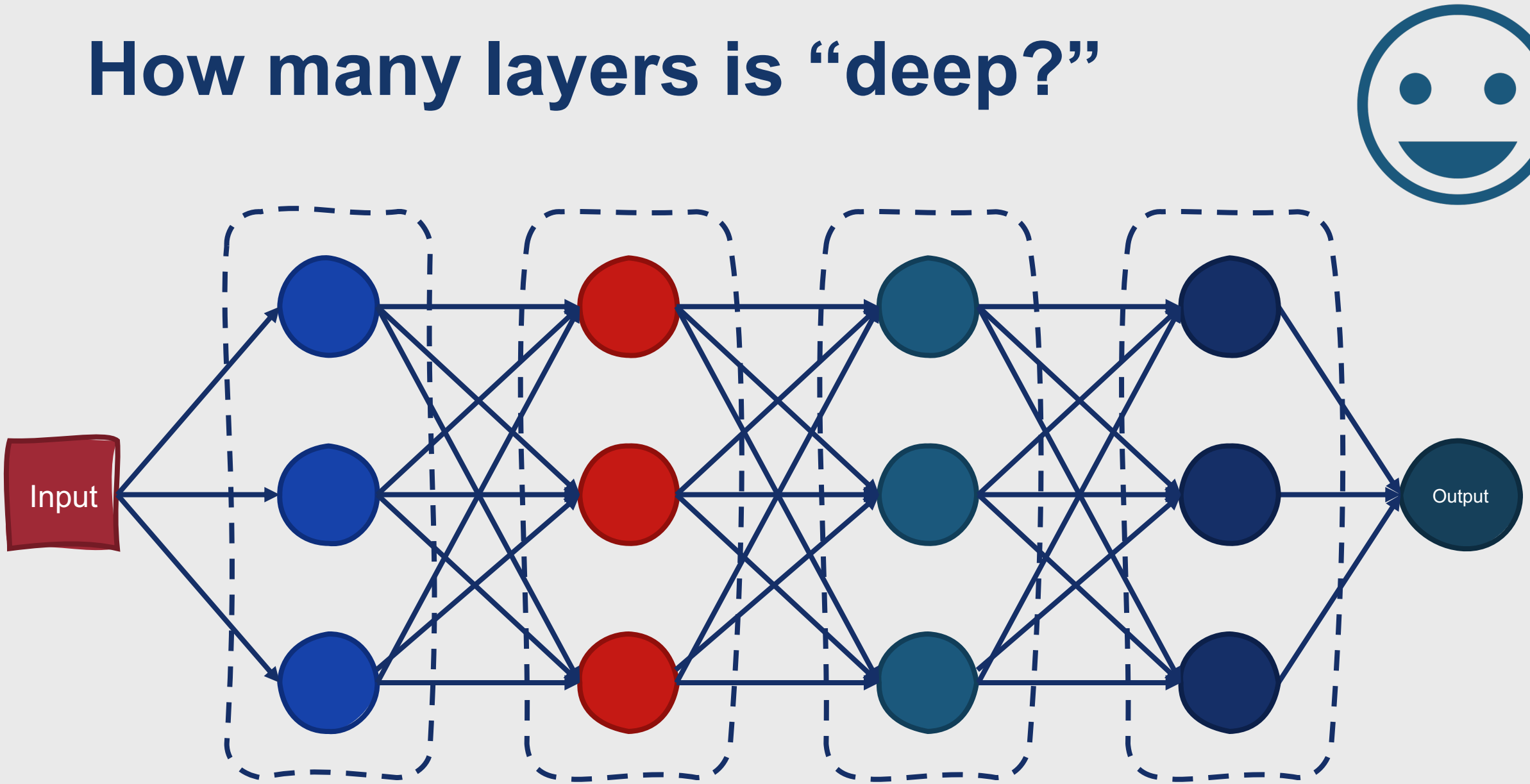
How many layers is “deep?”



How many layers is “deep?”

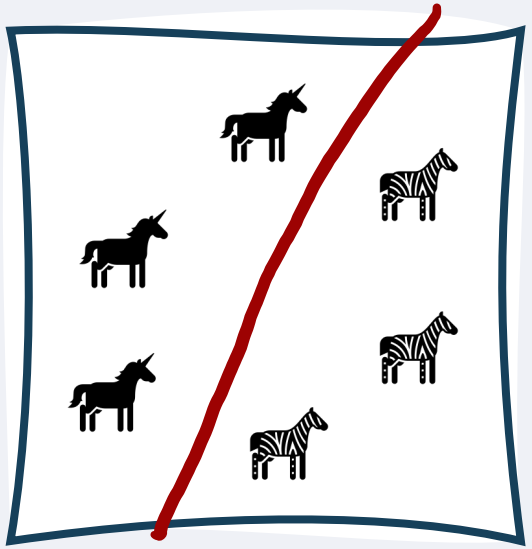


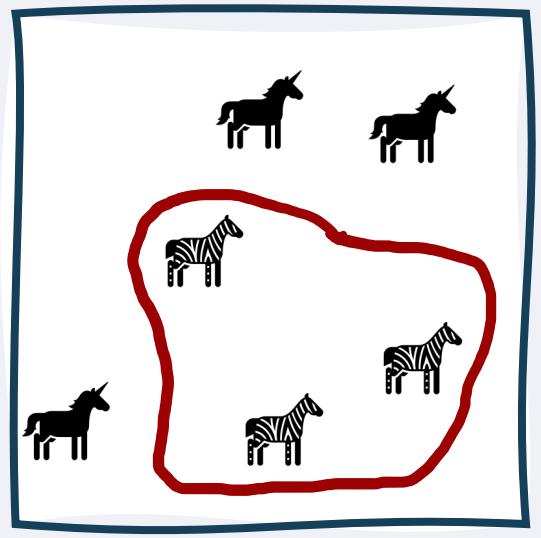
How many layers is “deep?”



Neural networks tend to be more powerful than traditional classification algorithms.

- Traditional classification algorithms usually assume that data is **linearly separable**
- In contrast, neural networks learn **nonlinear functions**





Neural networks also commonly use different types of features from traditional classification algorithms.

Traditional classification

- **Manually engineer** a set of features and extract them for each instance
  - Part-of-speech label
  - Number of exclamation marks
  - Sentiment score

Neural networks

- **Implicitly learn** features and extract those for each instance
  - Word embeddings

Neural networks aren't necessarily the best classifier for all tasks!

Learning features **implicitly** requires a lot of data

In general, deeper network → more data needed

Thus, neural nets tend to work very well for large-scale problems, but not that well for small-scale problems