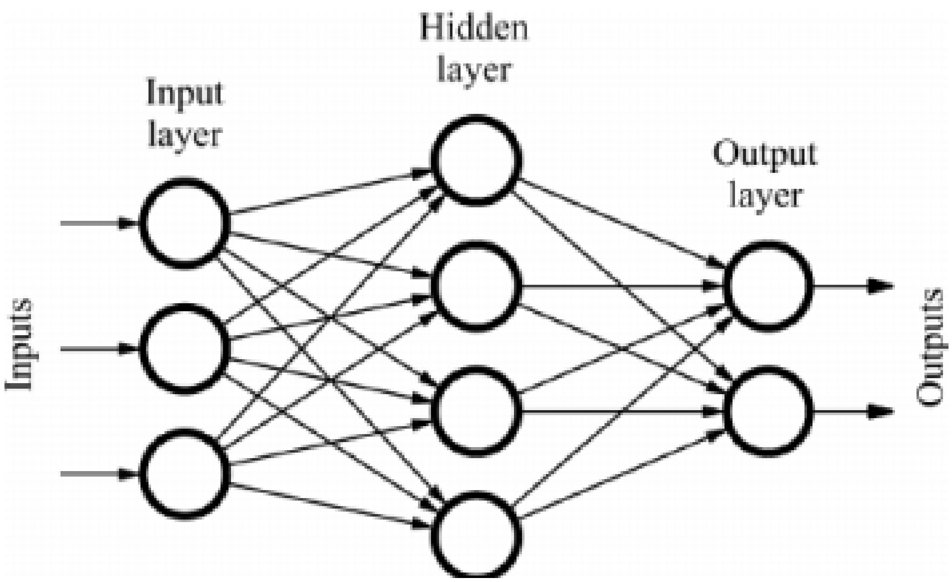
Introduction to Deep Learning

Feedforward Neural Networks:

A Feed Forward Neural Network is an artificial neural network in which the connections between nodes do not form a cycle. This is the most basic neural network architecture which is the building block for most other neural networks.



Why do we need Neural Networks?

As we have got more and more data, we want to make predictions from almost everything.

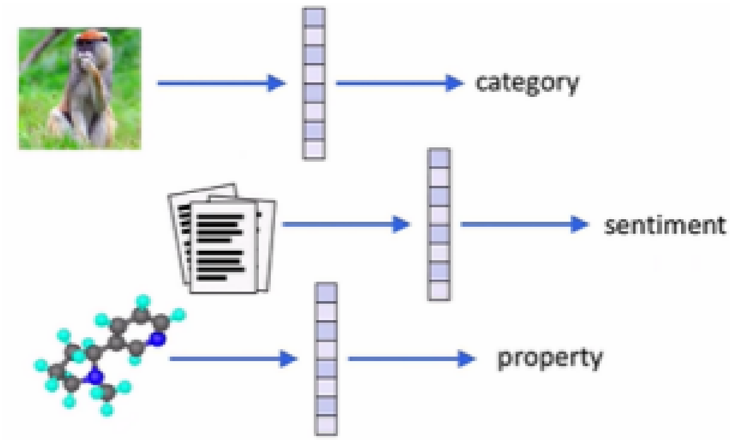
Example:

In images, we want computers to detect every object. In medical images, we want to recognize specific diseases.

We may want to make predictions from molecules for drug design, etc.

How can we learn from the data?

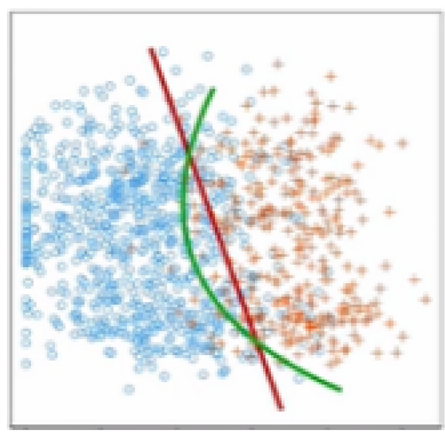
A general strategy is to encode the data as useful, informative feature vectors.



Once we have an encoding of the data that we have, then we can make predictions.

Recap of Feature Encodings that we have seen before:

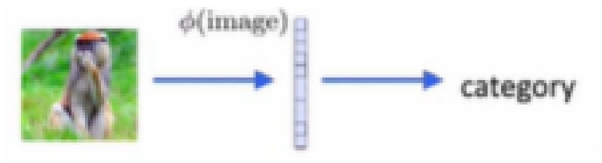
- Linear Classifiers: compare $\theta^T X$ to a threshold
  - Learn 'good' vector θ and threshold
- Nonlinear Classifiers: compare h(X) to a threshold
  - Learn 'good' function h and threshold



- Nonlinear classifier $\theta_1 X_1 + \theta_2 X_2 + \theta_{12} X_1 X_2$, is actually **linear** if we redefine $X = (X_1, X_2, X_1 X_2)$

- Feature-based linear classifier: compare $\theta^T\phi(X)$ to a threshold.

Here, we knew what we wanted, so we easily created a new representation of the data. But, for more general data, the encoding could be a bit more complicated. For example, finding good encodings for image data to classify them as some kind of animal is significantly harder than the previous example of separating different colors with a classifier.



And, Neural Networks learn these encodings **from the data**, instead of handcrafting the encodings like in previous examples.

The **challenge** here is: Now we have to learn encodings and predictions simultaneously. So, we need to train these feature encoding and classification stages together in one step. That's why Neural networks typically require more data.

Deep Learning: Reasons for success

- **Lots of data:** many problems can only be solved at scale.
- **Computational Resources(e.g GPUs):** we have access to systems that support deep ML algorithms at scale.
- **Large models are easier to train:** can be successfully estimated with simple gradient-based algorithms.
- **Flexible building blocks:** common representations, diversity of architecture choices.