10 Deep Learning

回顾：神经网络

Any continuous function f

$$f : R^N \to R^M$$

can be realized by an NN with one hidden layer.

(given **enough** hidden neurons)
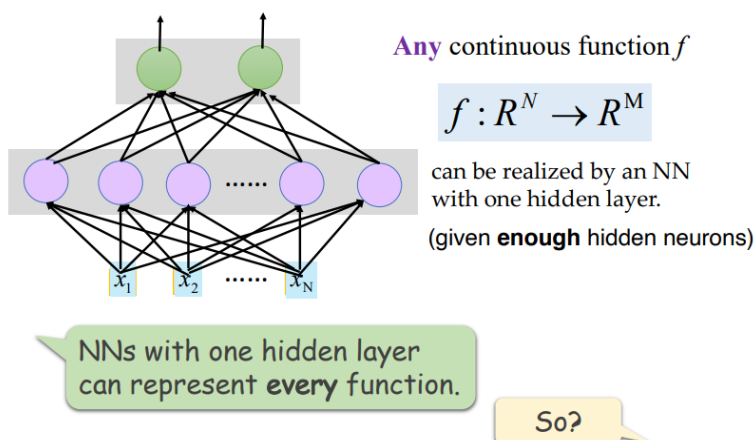
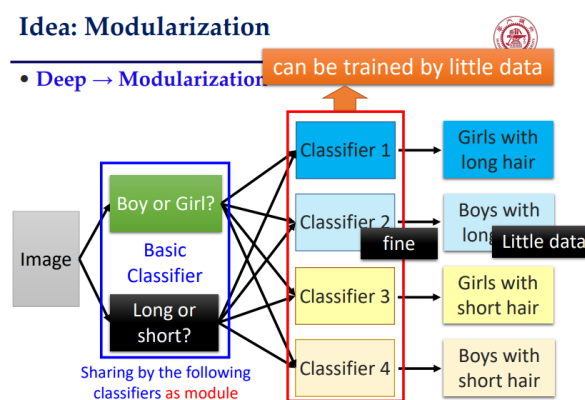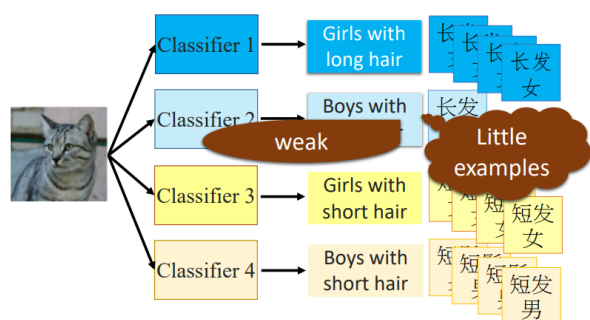NNs with one hidden layer can represent **every** function.

So?

The deeper, the better?

左下图要直接分成四类，第二种训练数据太少，所以导致训练结果不好。

右下图保证了basic classifier有足够的训练数量，正是因为在前一步划分后问题变得简单了，所以在中间的分类器用少量的数据训练就可以了。

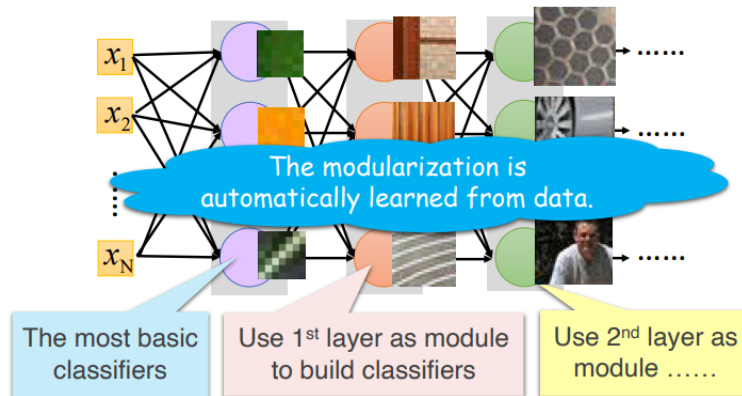Deep → Modularization（模块化）

E.G. How to express a picture to a trainable classifier? 层次化自动学习，第一层学习纹理，第二层学习小部件，第三层学习有含义的像素结构
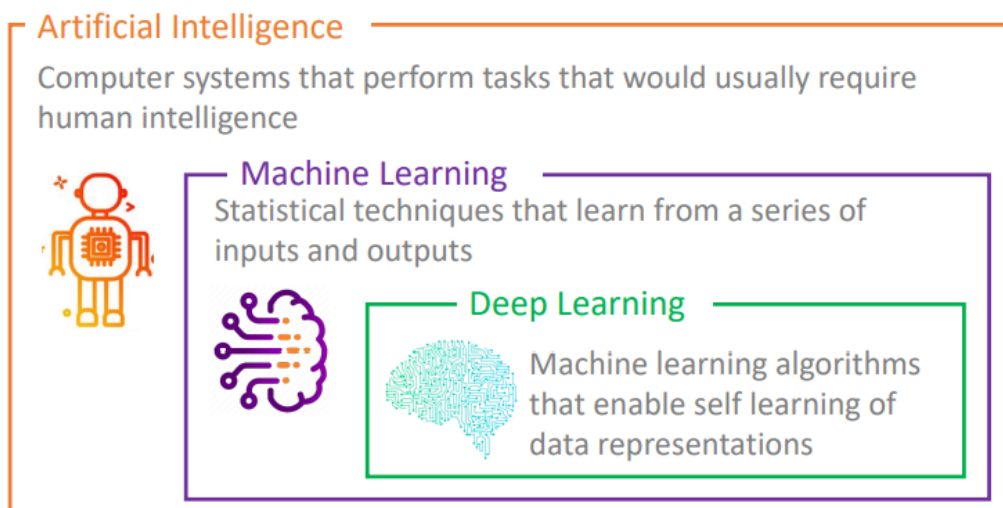
So, the deeper, the better!
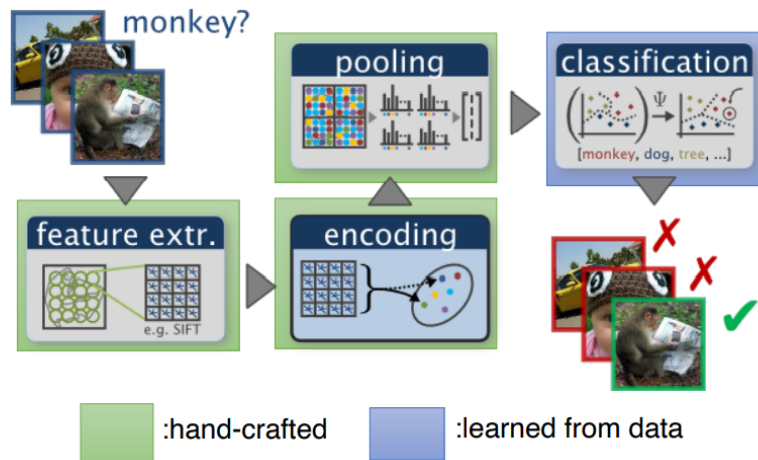
What is Deep Learning?

ML is a branch of machine learning based on a set of algorithms that attempt to model high-level abstractions in data by using model architectures, with complex structures or otherwise, composed of multiple non-linear transformations.

层次关系：人工智能→机器学习→深度学习

浅度学习人工的过程比较多，深度学习有trainable feature/trainable classifier

Why deep learning? Big Data / Big Machine / Big Model

神经网络：Loss Functions / Training

前馈神经网络(FNN) 1981

卷积神经网络(CNN) 1989

循环神经网络(RNN) 1986　　LSTM 1997

补充说明

深度学习的基础理论（如PLA/MLP/BP）其实在1986年前就已经基本成熟

深度学习在2006年后发展起来

因素：数据量增大、分布式流行、硬件发展 → 深度学习效果好