

# Neural Network Basics

March 5, 2019

## 1 What is Machine Learning?



Actually it's something like this.

Source:xkcd

## 2 Why Bother to learn?

Because...

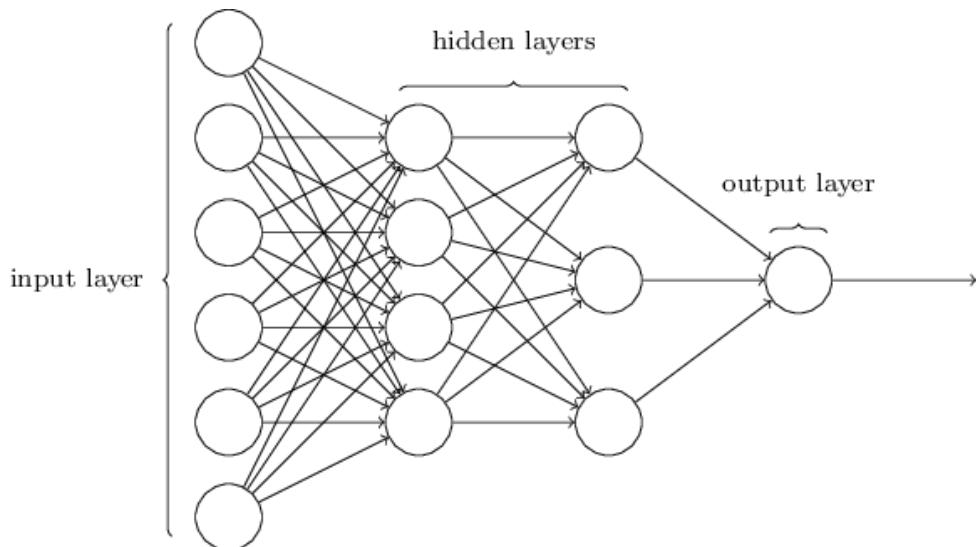


image.png



Source: Google Images

## 2.1 Neural Networks and Deep Learning

**What is a Neural Network?** Let's skip the definitions for which there's wikipedia at your disposal  
**For now I'll say they are universal function approximators** Have a look at this

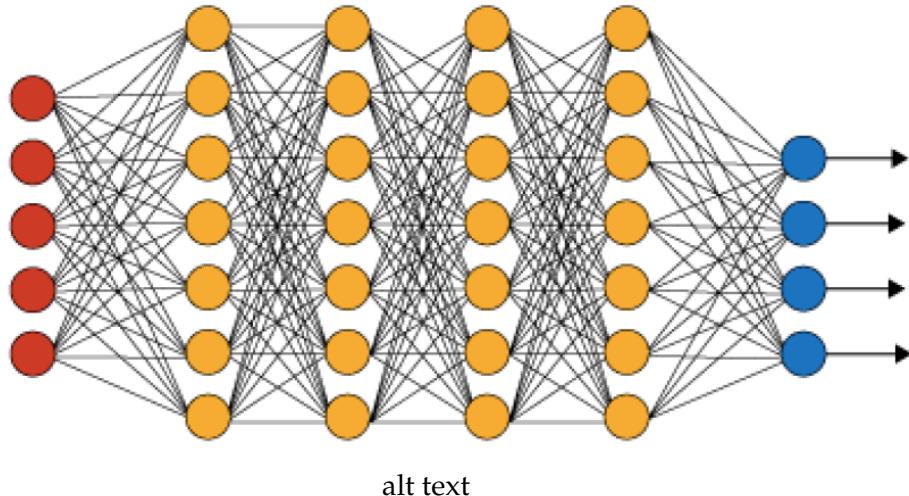
**2.1.1 Looks Familiar?**

**2.2 Lets go deep.**

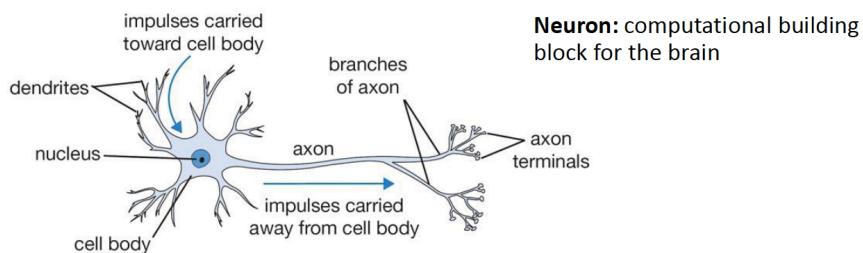
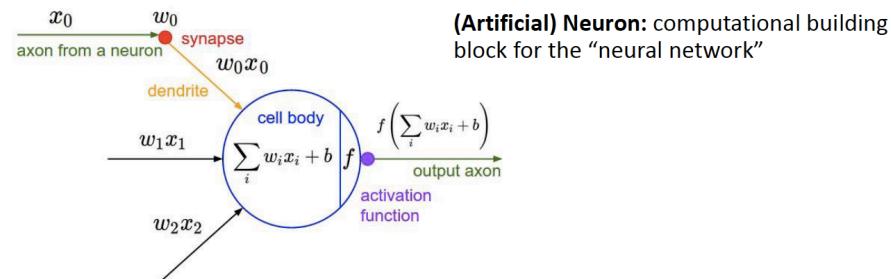
**2.3 Neuron: Biological Inspiration for Computation**

**Supervised Learning with Neural Network**

# Deep Learning Neural Network



alt text



biological\_inspiration.png

Input (X)	Output (y)	Application	
Home Features	Price	Real Estate	Standard Neural Networks
Ad, User info	Click on Ad? (0/1)	Online Advertising	Standard Neural Networks
Image	Object (1,2,...,100)	Photo Tagging	CNN
Audio	Text Transcript	Speech Recognition	RNN
English	Chinese	Machine Translation	RNN
Image, Radar Info	Position of other cars	Autonomous Driving	Custom/Hybrid

application.jpeg



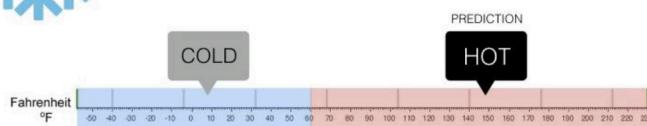
## Regression

What is the temperature going to be tomorrow?



## Classification

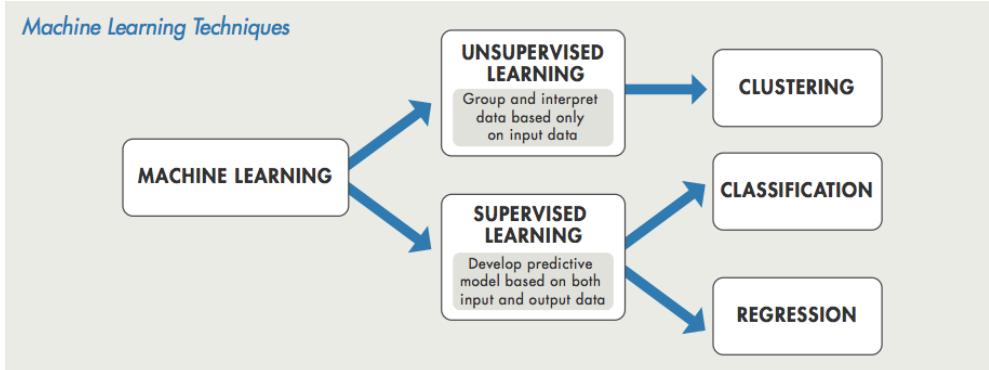
Will it be Cold or Hot tomorrow?



reg\_vs\_class.png

## 2.4 Regression vs Classification

Wait a minute, What did you say? Supervised, Unsupervised, What's that?



### Why is DeepLearning taking off?

The vertical axes of the diagram you can see the performance of an algorithm (e.g. it's prediction accuracy) and at the horizontal axes you can see the amount of data (Labelled Data).

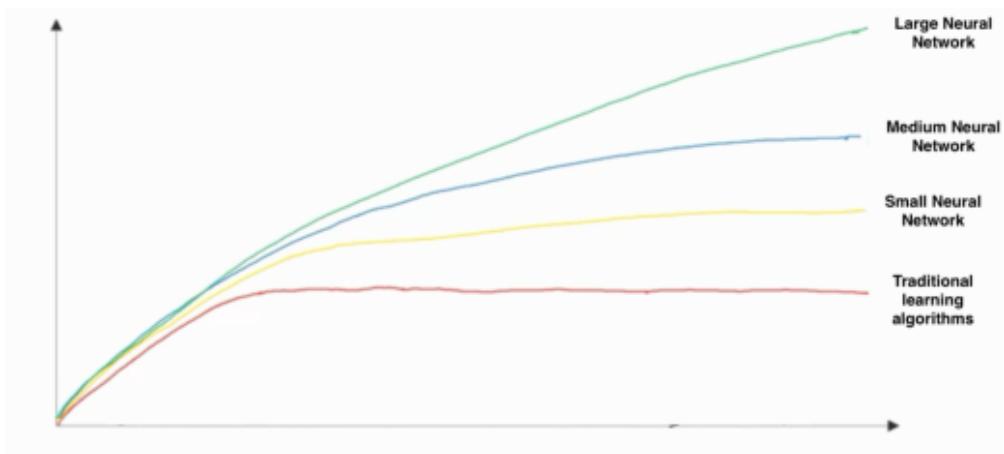
You can also see that the performance of traditional learning algorithms (logistic Regression, SVM's etc.) increases at the beginning with an increase of the amount of data but that it plateaus at a certain level and stops improving its performance.

The thing is that we have accumulated huge amounts of data over the last decades where our traditional learning algorithms can't take advantage of, which is where Deep Learning comes into play.

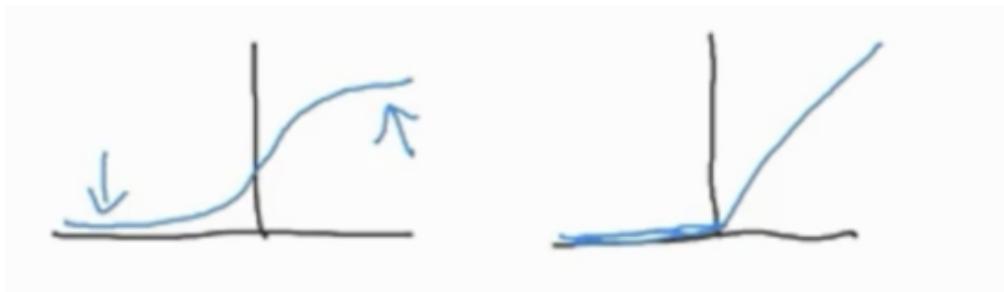
Large Neural Networks (e.g. Deep Learning) are getting better and better the more data you put into them. Andrew NG, a leading AI scientist, said that the three main forces which improve Neural Networks are:

1. Data
2. Computation
3. Algorithms

The recent breakthroughs in the development of algorithms are mostly due to making them



take\_off.png



relu\_sigmoid.png

run much faster than before, which makes it possible to use more and more data. For an example, a big advancement came from switching from a Sigmoid function (left picture) to a rectified-Linear-Unit function (right picture).

The other reason why fast computation is important is that, the below cycle must be faster.

**Bringing more data to a model is almost always beneficial.**

**Deep Learning approaches improve with more data**

## 2.5 Deep Learning Representation

## 2.6 Deep Learning

### 2.6.1 Aka. Here comes the exciting Part

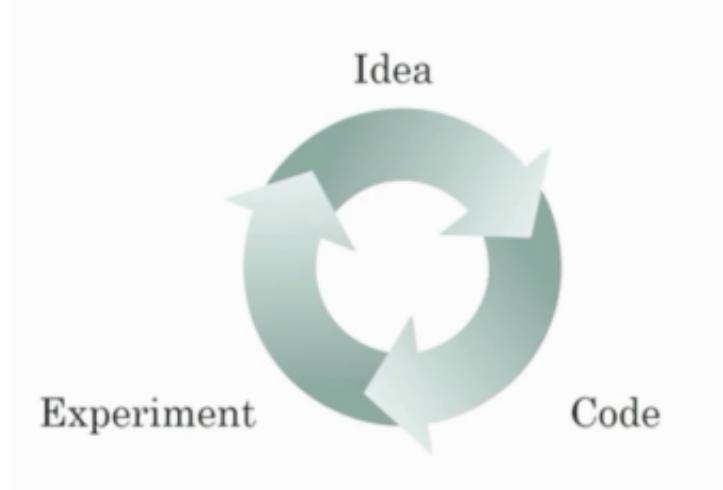
## 2.7 Deep Learning Tools

1. CUDA
2. Theano
3. Caffe2
4. Tensorflow 2.0
5. Pytorch 1.0

## 2.8 What is Tensorflow?

## 2.9 Deep Learning is Representation Learning

(aka Feature Learning)



basic\_cycle.png

## 2 Deep Learning representations

For representations:

- nodes represent inputs, activations or outputs
- edges represent weights or biases

Here are several examples of Standard deep learning representations

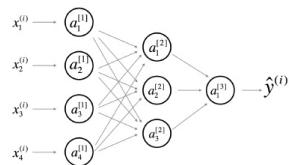


Figure 1: Comprehensive Network: representation commonly used for Neural Networks. For better aesthetic, we omitted the details on the parameters ( $w_{ij}^{[l]}$  and  $b_i^{[l]}$  etc...) that should appear on the edges

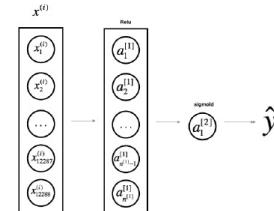


Figure 2: Simplified Network: a simpler representation of a two layer neural network, both are equivalent.

Representation.jpeg

- **What is it:**  
Extract useful patterns from data.
- **How:**  
Neural network + optimization
- **How (Practical):**  
Python + TensorFlow & friends
- **Hard Part:**  
Good Questions + Good Data
- **Why now:**  
Data, hardware, community, tools, investment
- **Where do we stand?**  
Most big questions of intelligence have not been answered nor properly formulated

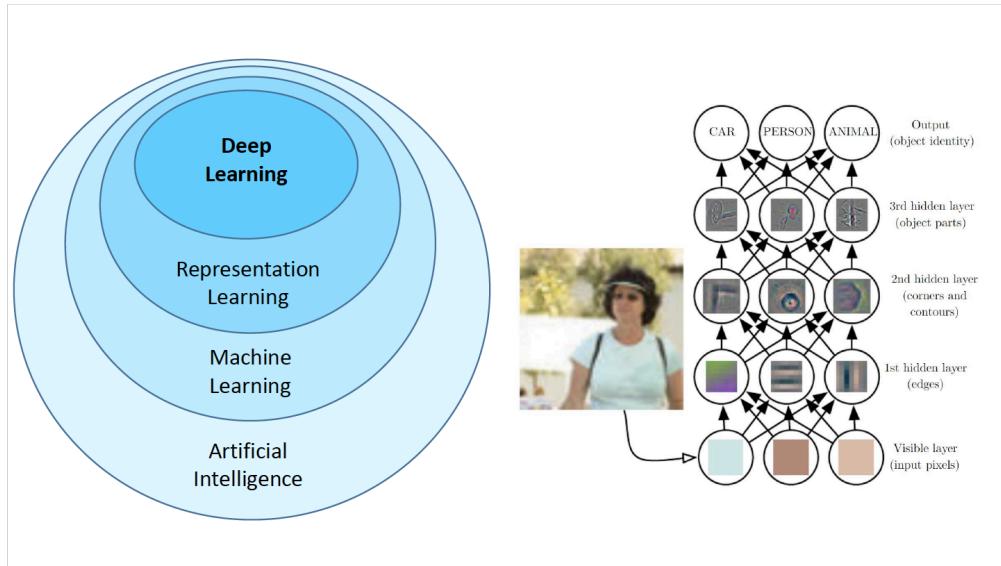
- Exciting progress:**
- Face recognition
  - Image classification
  - Speech recognition
  - Text-to-speech generation
  - Handwriting transcription
  - Machine translation
  - Medical diagnosis
  - Cars: drivable area, lane keeping
  - Digital assistants
  - Ads, search, social recommendations
  - Game playing with deep RL

DeepLearning.png

- **What is it:** Deep Learning Library (*and more*)
  - **Facts:** Open Source, Python, Google
- **Community:**
  - 117,000+ GitHub stars
  - TensorFlow.org: Blogs, Documentation, DevSummit, YouTube talks
- **Ecosystem:**
  - **Keras:** high-level API
  - **TensorFlow.js:** in the browser
  - **TensorFlow Lite:** on the phone
  - **Colaboratory:** in the cloud
  - **TPU:** optimized hardware
  - **TensorBoard:** visualization
  - **TensorFlow Hub:** graph modules
- **Alternatives:** PyTorch, MXNet, CNTK

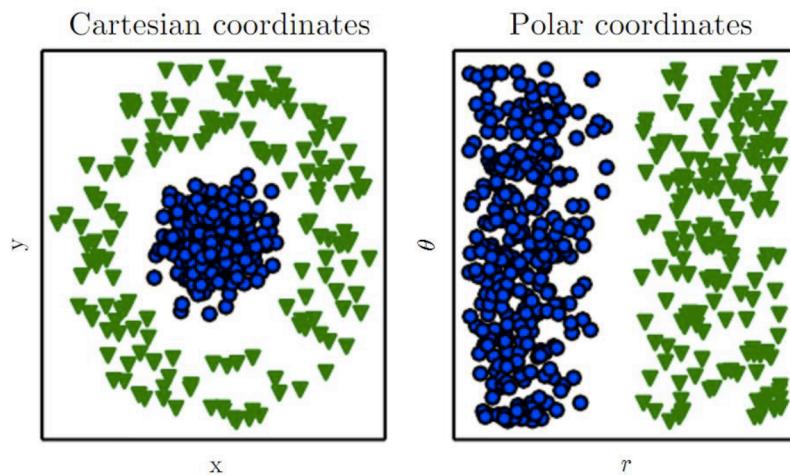
- Extras:**
- Swift for TensorFlow
  - TensorFlow Serving
  - TensorFlow Extended (TFX)
  - TensorFlow Probability
  - Tensor2Tensor

tensorflow.png

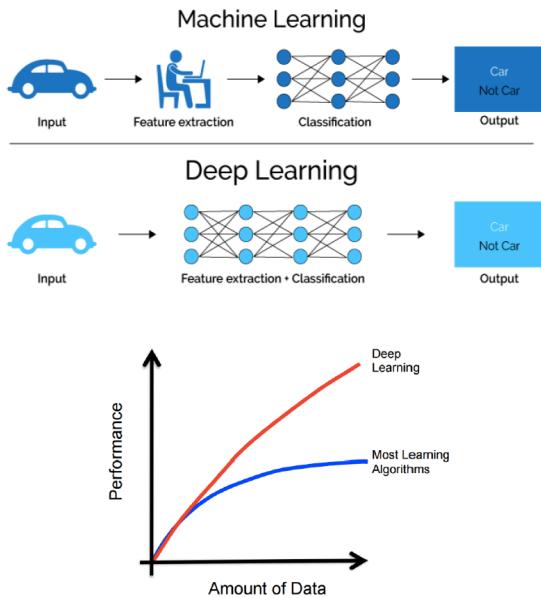


Representation\_learning.png

## Representation Matters



Representation\_matters.png



why\_dl?.png

## 2.10 WHY DEEP LEARNING?

### 2.11 What is a Neural Network?

By the Example of House Price Prediction

### 2.12 DeepLearning for human and machine

### 2.13 Data Augmentation

### 2.14 Deep Learning: Training and Testing

### 2.15 Neuron : Forward pass

### 2.16 Steps Involved in forward propagation!!

### 2.17 How Neural Networks Learns : BackPropagation of Errors

#### 2.17.1 Steps Involved

Input  $x$ " role="presentation" style="position: relative;">xx

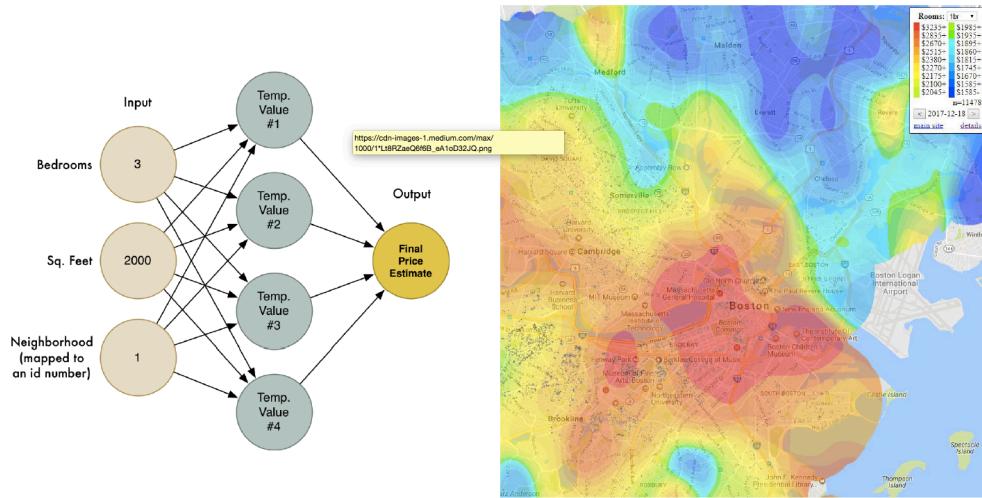
: Set the corresponding activation  $a_1$ " role="presentation" style="position: relative;">a1a1  
for the input layer.

Feedforward: For each  $l=2,3,\dots,L$ " role="presentation" style="position: relative;">l=2,3,...,L  
compute  $z_l=w_{l-1}a_{l-1}+b_l$ " role="presentation" style="position: relative;">z<sub>l</sub>=w<sub>l-1</sub>a<sub>l-1</sub>+b<sub>l</sub>

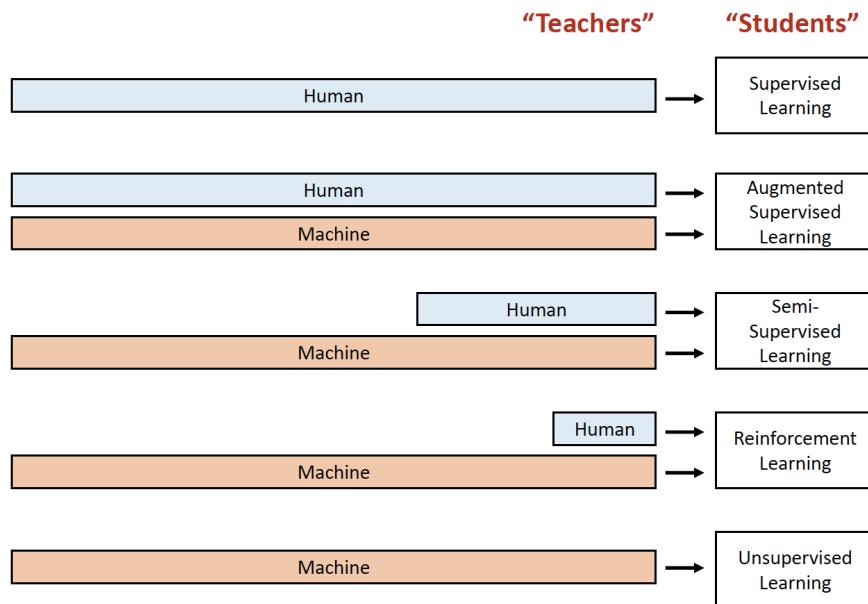
$z_l=w_{l-1}a_{l-1}+b_l$ " role="presentation" style="position: relative;">z<sub>l</sub>=w<sub>l-1</sub>a<sub>l-1</sub>+b<sub>l</sub>  
and  $a_l=\sigma(z_l)$ " role="presentation" style="position: relative;">a<sub>l</sub>=σ(z<sub>l</sub>)

Output error  $\delta_L=(y-L)$ " role="presentation" style="position: relative;">δ<sub>L</sub>=(y-L)

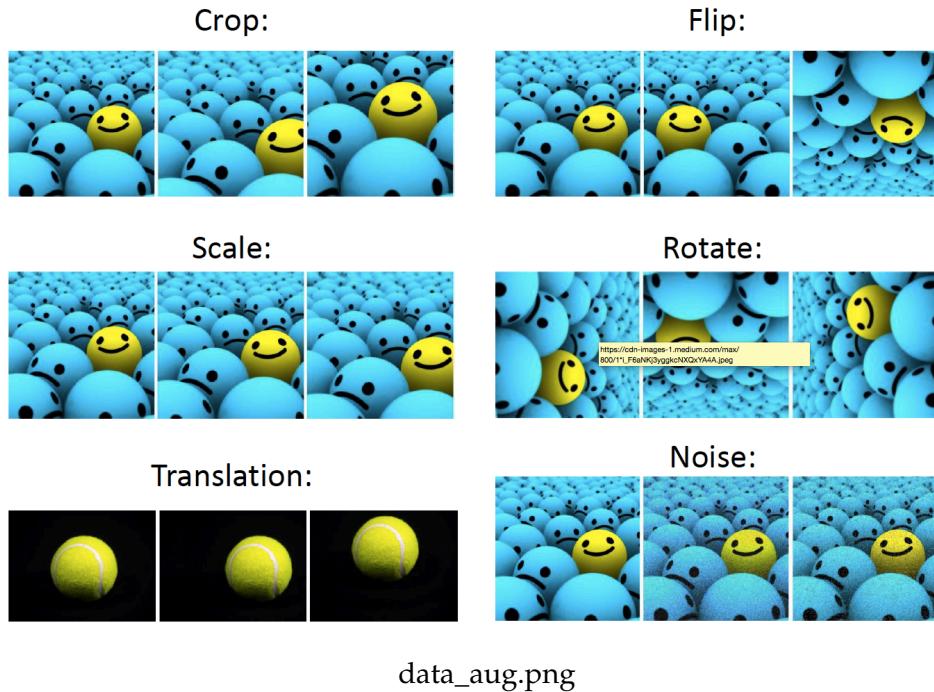
Output error  $\delta_L=(y-L)$ " role="presentation" style="position: relative;">δ<sub>L</sub>=(y-L)



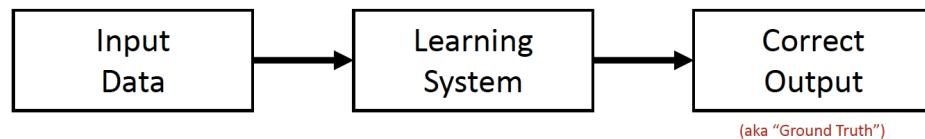
house\_pred\_exmaple.png



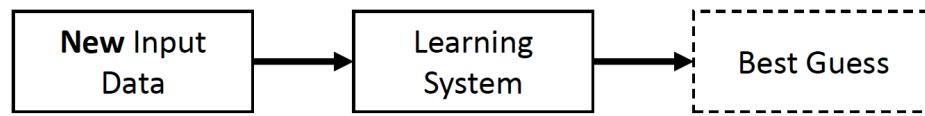
dl\_human-machine.png



### Training Stage:



### Testing Stage:



and testing.bb

---

training%20and%20testing.png