You have already learned the Machine Learning. When?

?

You have already learned the Machine Learning. When?

$$2x = 6 \tag{1}$$

$$(2x - 6) = 0 \tag{2}$$

$$x = ? \tag{3}$$

You have already learned the Machine Learning. When?

$$2x = 6 \tag{4}$$

$$(2x - 6) = 0 \tag{5}$$

$$x = ? \tag{6}$$

$$x = 6/2 = 3 \tag{7}$$

$$2(3) - 6 = 0 \tag{8}$$

You have already learned the Machine Learning. When?

$$2a + b + c = 4 \tag{9}$$

$$a + 3b + 2c = 5 \tag{10}$$

$$a = 6 \tag{11}$$

You have already learned the Machine Learning. When?

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \tag{12}$$

You have already learned the Machine Learning. When?

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 0 & 0 \end{bmatrix}, x = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \quad (13)$$

You have already learned the Machine Learning. When?

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \tag{14}$$

$$Ax = b \tag{15}$$

$$(Ax - b) = ? \tag{16}$$

You have already learned the Machine Learning. When?

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \tag{17}$$

$$Ax = b \tag{18}$$

$$(Ax - b) = 0 \tag{19}$$

$$x = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = ? \tag{20}$$

What is Regression

Regression?

What is Regression

$$x + y = z$$

Solving Ax=b

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \tag{21}$$

$$X\ W = Y \tag{22}$$

$$(X\ W - Y) = 0 \tag{23}$$

$$W = \begin{bmatrix} a \\ b \\ c \end{bmatrix} =? \tag{24}$$

$$X^{\dagger}\ X\ W = X^{\dagger}\ Y \tag{25}$$

$$I\ W = X^{\dagger}\ Y \tag{26}$$

$$W = X^{\dagger}\ Y \tag{27}$$

Decimal Value Prediction

ID	digit1	digit2	digit3	value
1	0	0	0	**0**
2	0	0	1	**1**
3	0	1	0	**2**
4	0	1	1	**3**
5	1	0	0	**4**
6	1	0	1	**5**
7	1	1	0	**6**
8	1	1	1	**7**

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} w1 & w2 & w3 \end{bmatrix}^T = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{bmatrix} \tag{28}$$

$$\mathbf{X}\ \mathbf{w}^T = \mathbf{y} \tag{29}$$

$$\mathbf{X}^{\dagger}\ X\ \mathbf{w}^T = X^{\dagger}\ \mathbf{y} \tag{30}$$

$$\mathbf{w}^T = X^{\dagger}\ \mathbf{y} \tag{31}$$

$$\mathbf{w}^T = \mathbf{X}^\dagger \, \mathbf{y} = \mathbf{X}^\dagger \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 1 \end{bmatrix}^T = \begin{bmatrix} w1 & w2 & w3 \end{bmatrix}^T \quad (32)$$

$$\mathbf{X}\ \mathbf{w}^T = \mathbf{y} \tag{33}$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 & 1 \end{bmatrix}^T = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{bmatrix} \tag{34}$$

$$\mathbf{X}\ \mathbf{w}^T = \mathbf{y}^{pre} \tag{35}$$

$$cost\ \ function = \frac{1}{m}\sum_{i}^{m}(\mathbf{y}_i - \mathbf{y}_i^{pre})^2 \tag{36}$$

$$training\ \ error = sum(abs(\mathbf{y} - \mathbf{y}^{pre})) \tag{37}$$

$$\mathbf{y} - \mathbf{y}^{pre} = \begin{bmatrix}0\\1\\2\\3\\4\\5\\6\\7\end{bmatrix} - \begin{bmatrix}0\\1\\2\\3\\4\\5\\6\\7\end{bmatrix} = sum\begin{bmatrix}0\\0\\0\\0\\0\\0\\0\\0\end{bmatrix} = 0 \tag{38}$$

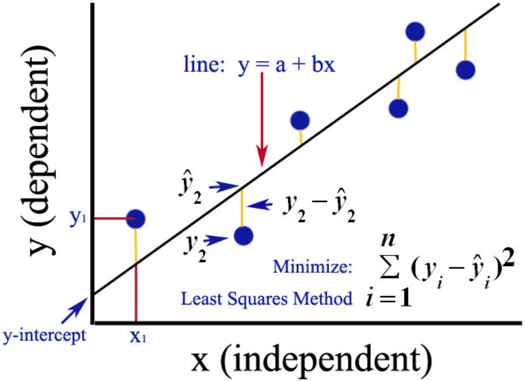
$$\begin{bmatrix} digit1 & digit2 & digit3 \end{bmatrix} \begin{bmatrix} w1 & w2 & w3 \end{bmatrix}^T = [value] \tag{39}$$

$$\mathbf{digit1 * w1 + digit2 * w2 + digit3 * w3 = value} \tag{40}$$

$$\begin{bmatrix} digit1 & digit2 & digit1 \end{bmatrix} \begin{bmatrix} 4 & 2 & 1 \end{bmatrix}^T = [value] \tag{41}$$

$$\mathbf{digit1 * 4 + digit2 * 2 + digit3 * 1 = value} \tag{42}$$

Linear Regression



source: solutions4statistics.com

Decimal Value Prediction

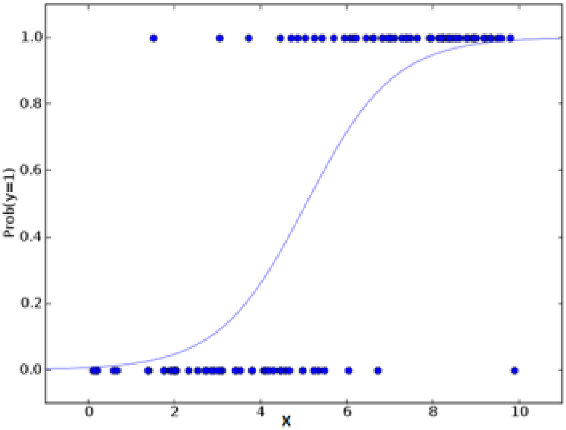
ID	digit1	digit2	digit3	value	decision
1	0	0	0	**0**	**0**
2	0	0	1	**1**	**0**
3	0	1	0	**2**	**0**
4	0	1	1	**3**	**0**
5	1	0	0	**4**	**1**
6	1	0	1	**5**	**1**
7	1	1	0	**6**	**1**
8	1	1	1	**7**	**1**

$$\begin{bmatrix} digit1 & digit2 & digit3 \end{bmatrix} \begin{bmatrix} w1 & w2 & w3 \end{bmatrix}^T = [value] \tag{43}$$

$$\mathbf{digit1 * w1 + digit2 * w2 + digit3 * w3 = value} \tag{44}$$

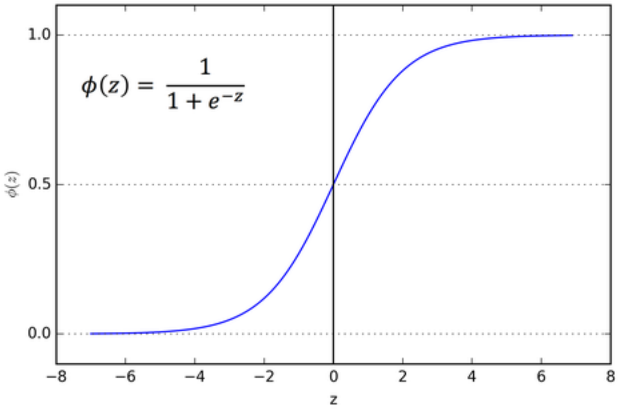
$$Prediction = \begin{cases} 1 & if\ 4 \geqslant value \\ 0 & else \end{cases} \tag{45}$$

Logistic Regression



source: `solutions4statistics.com`

Logistic - Sigmoid Function



https://sebastianraschka.com/images/faq/logisticregr-neuralnet/sigmoid.png

Logistic - Sigmoid

$$\Phi(z) = \frac{1}{1 + exp^{-z}} \tag{46}$$

$$\Phi(-6) = \frac{1}{1 + exp^{-(-6)}} = \frac{1}{1 + 403.42} = 0.0024 \tag{47}$$

$$\Phi(0) = \frac{1}{1 + exp^{0}} = \frac{1}{1 + 1} = 0.5 \tag{48}$$

$$\Phi(6) = \frac{1}{1 + exp^{-(6)}} = \frac{1}{1 + 0.0024} = 0.997 \tag{49}$$

Logistic Regression

$$\Phi(z) = \frac{1}{1 + exp^{-z}} \tag{50}$$

$$\mathbf{y} = \Phi(\mathbf{X}\ \mathbf{w}^T) = \frac{1}{1 + exp^{-(\mathbf{X}\ \mathbf{w}^T)}} \tag{51}$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} w1 & w2 & w3 \end{bmatrix}^T = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \tag{52}$$

$$\mathbf{X}\ \mathbf{w}^T = \mathbf{y} \tag{53}$$

$$\mathbf{X}^\dagger\ \mathbf{X}\ \mathbf{w}^T = \mathbf{X}^\dagger\ \mathbf{y} \tag{54}$$

$$\mathbf{w}^T = \mathbf{X}^\dagger\ \mathbf{y} \tag{55}$$

$$\mathbf{w}^T = \mathbf{X}^\dagger\,\mathbf{y} = \mathbf{X}^\dagger \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.24 & -0.11 & -0.11 \end{bmatrix}^T = \begin{bmatrix} w1 & w2 & w3 \end{bmatrix}^T \tag{56}$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1.24 & -0.11 & -0.11 \end{bmatrix}^T = sigmoid\left(\begin{bmatrix} 0.0 \\ -0.11 \\ -0.11 \\ -0.22 \\ 1.24 \\ 1.12 \\ 1.12 \\ 1.01 \end{bmatrix} \right)$$

$$\mathbf{X}\ \mathbf{w}^T = sigmoid(\mathbf{y}^{pre})$$

$$sigmoid\left(\begin{bmatrix} 0.0 \\ -0.11269202 \\ -0.11269202 \\ -0.22538404 \\ 1.24054754 \\ 1.12785552 \\ 1.12785552 \\ 1.0151635 \end{bmatrix}\right) = \begin{bmatrix} 0.5 \\ 0.47185 \\ 0.47185 \\ 0.44389 \\ 0.77565 \\ 0.75544 \\ 0.75544 \\ 0.73402 \end{bmatrix} \tag{57}$$

$$X\ \mathbf{w} = sigmoid(\mathbf{y}^{pre}) \tag{58}$$

$$cost\ \ function = \frac{1}{m}\sum_{i}^{m}(\mathbf{y}_i - sigmoid(\mathbf{y}_i^{pre}))^2 \tag{59}$$

$$training\ \ error = sum(abs(\mathbf{y} - sigmoid(\mathbf{y}^{pre})) \tag{60}$$

$$abs(\mathbf{y} - sigmoid(\mathbf{y}^{pre}) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{61}$$

$$training\ \ error = 0 \tag{62}$$

$$\begin{bmatrix} 0.5 \\ 0.47185 \\ 0.47185 \\ 0.44389 \\ 0.77565 \\ 0.75544 \\ 0.75544 \\ 0.73402 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \tag{63}$$

$$Prediction = \begin{cases} 1 & if \;\; sigmoid(\mathbf{y}^{pre}) \geqslant 0.5 \\ 0 & else \end{cases} \tag{64}$$

$$\begin{bmatrix} digit1 & digit2 & digit3 \end{bmatrix} \begin{bmatrix} 1.24 & -0.11 & -0.11 \end{bmatrix}^2 = [value] \tag{65}$$

$$\mathbf{digit1 * w1 + digit2 * w2 + digit3 * w3 = value} \tag{66}$$

$$sigmoid(\mathbf{y}^{pre}) = \frac{\mathbf{1}}{\mathbf{1 + exp^{-(digit1 * w1 + digit2 * w2 + digit3 * w3)}}} \tag{67}$$

$$Prediction = \begin{cases} 1 & if \;\; sigmoid(\mathbf{y}^{pre}) \geqslant 0.5 \\ 0 & else \end{cases} \tag{68}$$

Evaluating the model

- Accuracy

$$Accuracy = \frac{\#\ correctly\ \ classified\ \ instances}{total\#\ \ instances} \tag{69}$$

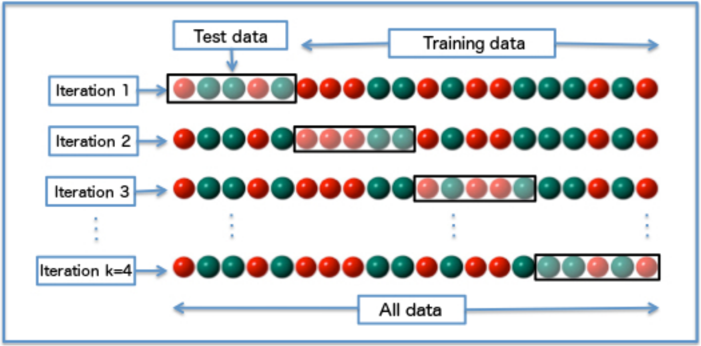
Evaluating the model

$$\begin{bmatrix}1\\0\\0\\0\\1\\1\\1\\0\end{bmatrix} == \begin{bmatrix}0\\0\\0\\0\\1\\1\\1\\1\end{bmatrix} = \begin{bmatrix}False\\True\\True\\True\\True\\True\\True\\False\end{bmatrix} \tag{70}$$

Accuracy = 6 / 8 * 100 = 75 %

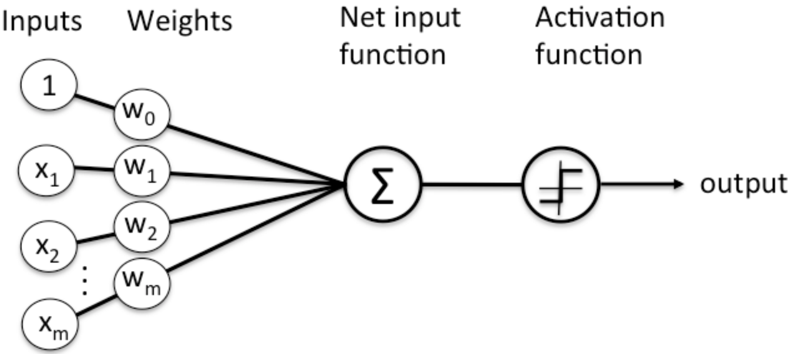
Improving the performance

10 - fold 10-cross validation



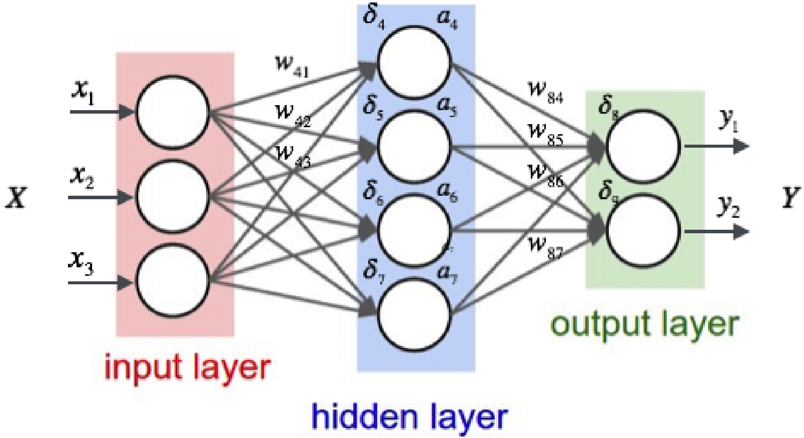
Source: wikipedia

Logistic Regression as a Neuron



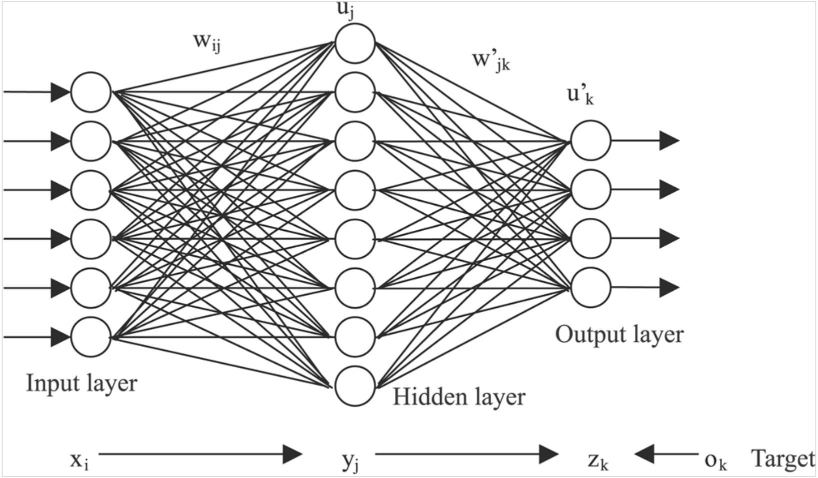
www.techmaru.com/technology/artificial-neural-networks/neural-network-elements

Neuron to Neurons



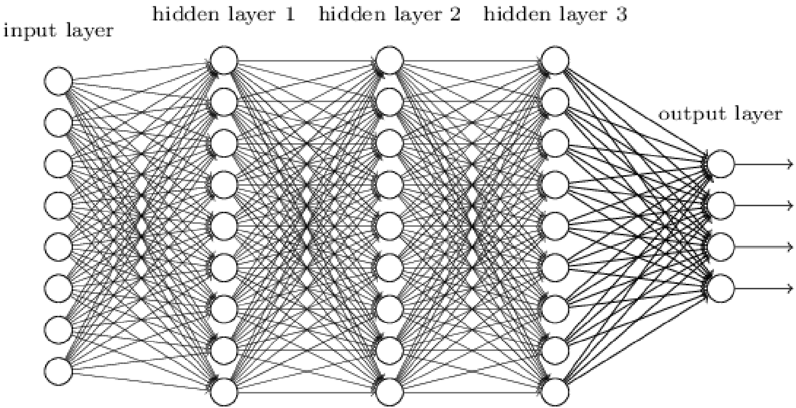
medium.com/@curiousily/tensorflow-for-hackers-part-iv-neural-network-from-scratch-1a4f504dfa8

Single Layer Network



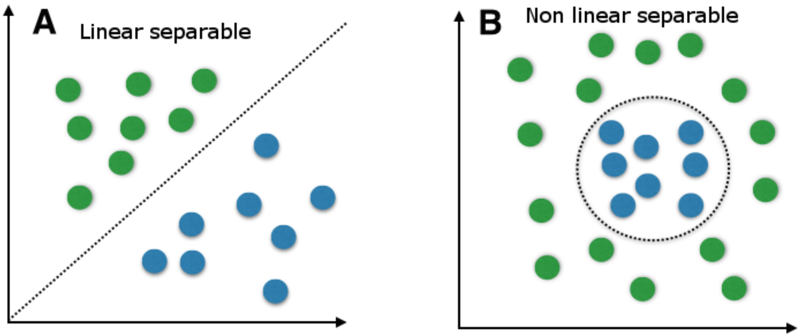
www.extremetech.com/extreme/215170-artificial-neural-networks-are-changing-the-world-what-are-they

Multi Layer Network



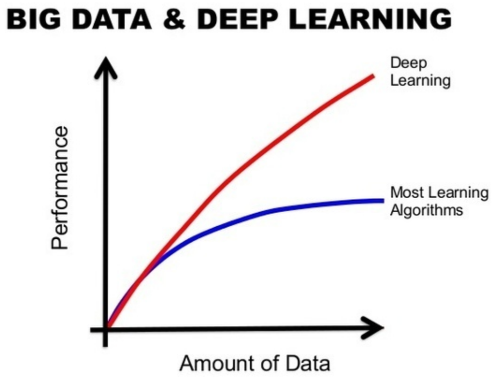
in.mathworks.com/matlabcentral/fileexchange/64247-simple-neural-network

Why Deep Learning?



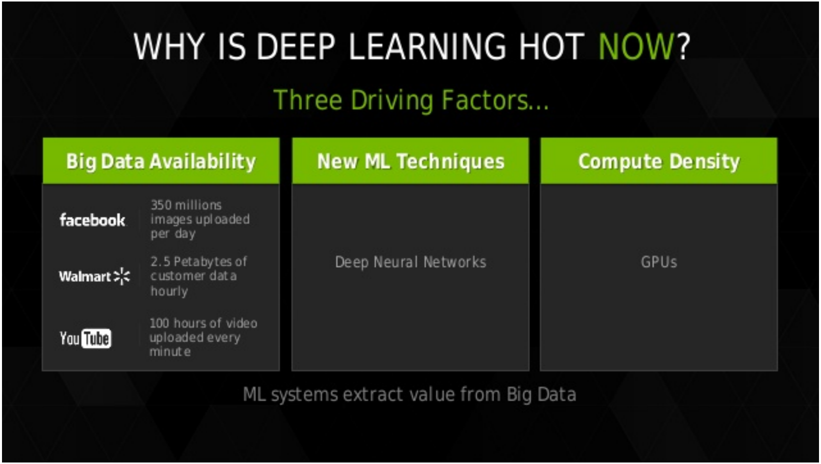
Source: https://leonardoaraujosantos.gitbooks.io/artificial-inteligence/content/linear_classification.html

Why Deep Learning?



Source: `https://qph.ec.quoracdn.net/main-qimg-bf69c291005e68620a1bef39ae8f029e-c`

Why now Deep Learning?



https://www.slideshare.net/DataScienceMD/deep-learning-with-gpus