

Session 1&2

Feed Forward Neural Networks (FFNN)

Deep Learning | Zahra Amini

Telegram: @zahraamini_ai & Instagram:@zahraamini_ai & LinkedIn: @zahraamini-ai

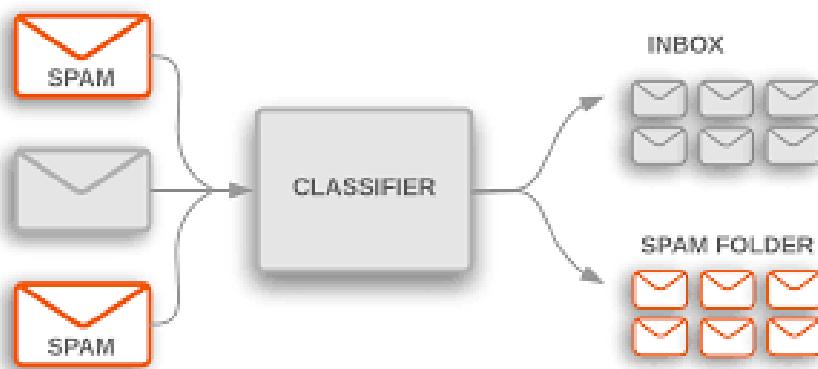
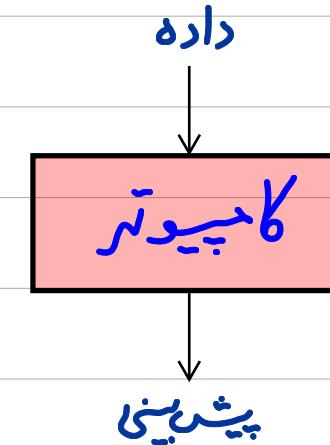
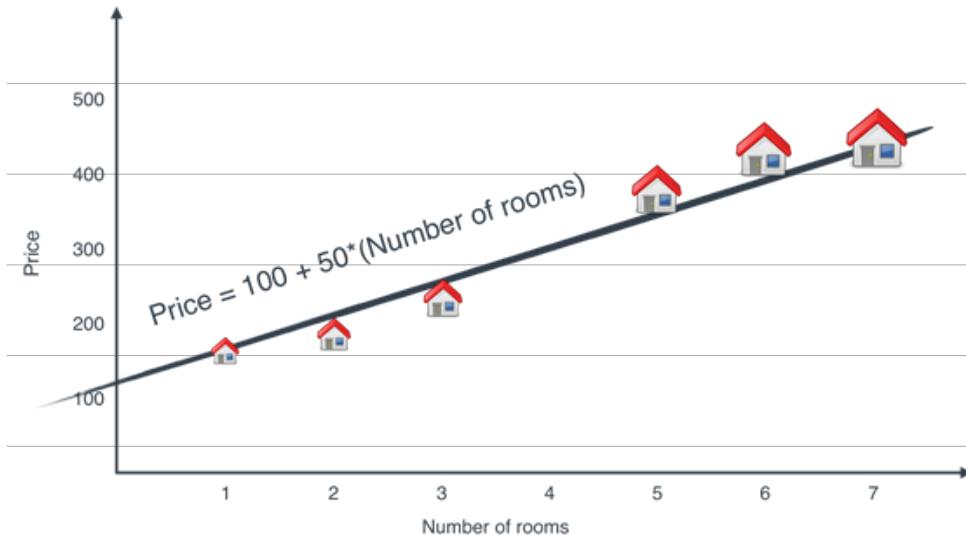
<https://zil.ink/zahraamini>

هرش مفهومی چیست ؟

به هر تکنیکی که کامپیوترها قادری سازد تا انسان رفتار کند.

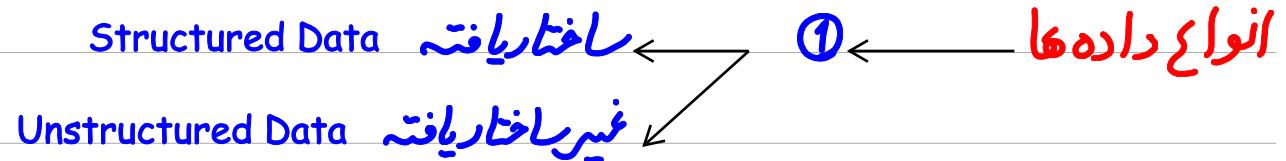
یادگیری ماشین چیست ؟

قابلیت یادگیری بدون نیاز به برنامه ریزی مستقیم یا صریح (if)

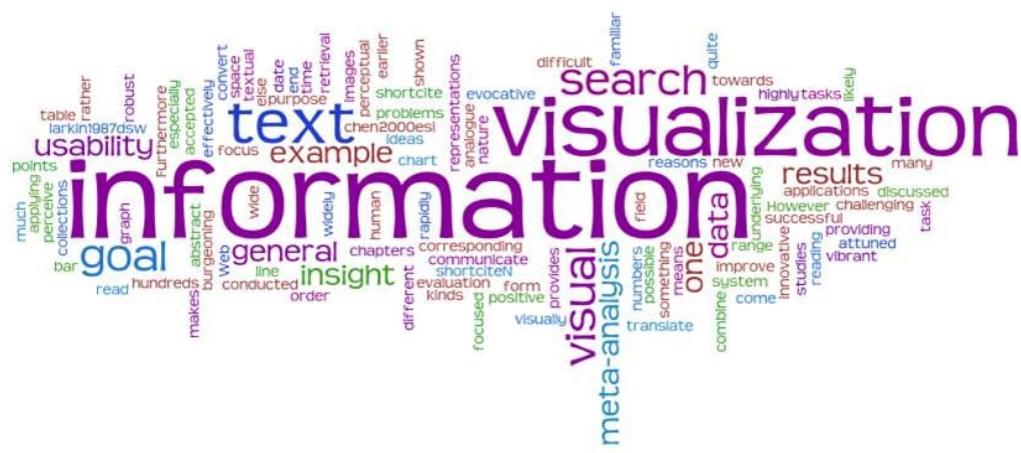


دیتابیس چیست؟

مجموعه داده‌های جمیع آوری شده برای آموزش و ارزیابی مدل.



ویژگی	داده‌های ساختار یافته	داده‌های غیرساختار یافته
ساختار	سازمان یافته و قابل پیش‌بینی	سازمان‌دهی نشده یا دارای ساختار نامنظم
ذخیره‌سازی	در پایگاه‌های داده رابطه‌ای (SQL)	در سیستم‌های پایگاه داده NoSQL یا به صورت فایل
مدیریت	آسان برای مدیریت به دلیل یکپارچگی و فهرست‌بندی	مدیریت دشوارتر به دلیل عدم وجود ساختار مشخص
پردازش	پردازش سریع و ساده به خاطر ساختار مشخص	نیاز به ابزارهای پیشرفته برای تحلیل و پردازش
تحلیل داده‌ها	تحلیل ساده و مستقیم	تحلیل پیچیده و گاهی اوقات نیازمند پردازش خاص
مثال‌ها	داده‌های مشتری، تراکنش‌های بانکی	ایمیل‌ها، فایل‌های ویدئویی، پست‌های شبکه‌های اجتماعی

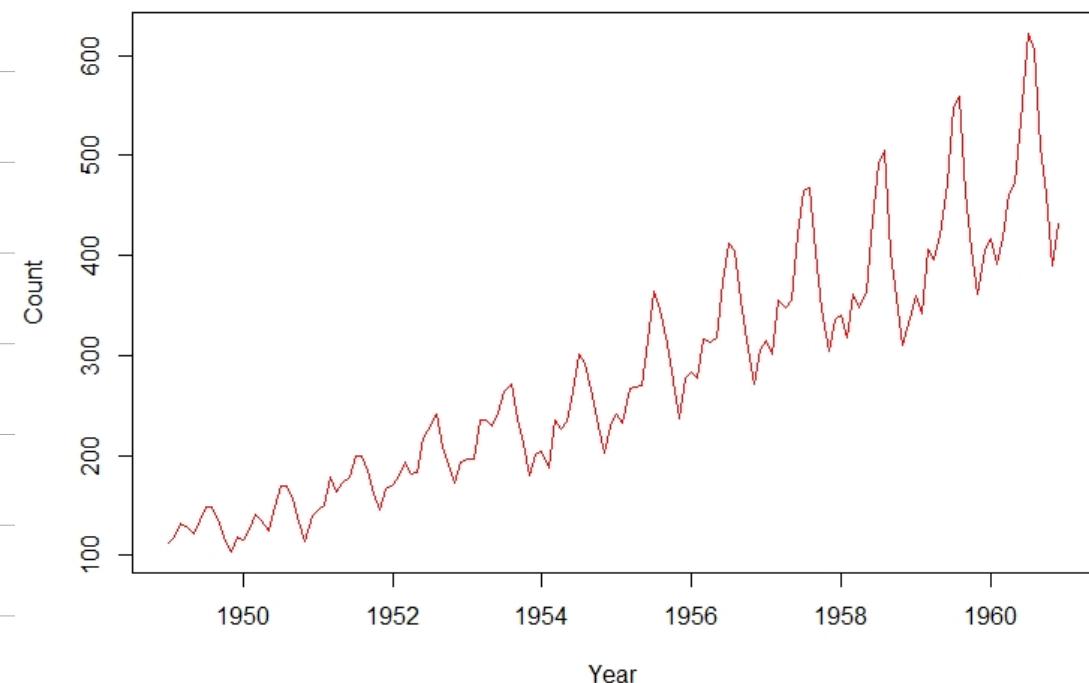


Tabular Data

ID	TOTAL ACTIONS	ACTION 1	ACTION 2	TOTAL TIME
10	120	80	40	0:50:05
11	255	130	125	1:40:03
12	180	100	80	1:20:19
13	305	205	100	1:58:58
14	71	50	21	0:35:41
15	418	310	108	2:08:18
16	222	150	72	1:32:58

Time series data

Monthly Airline Passenger Numbers



انواع داده‌ها

Categorical Data کتگوریکال

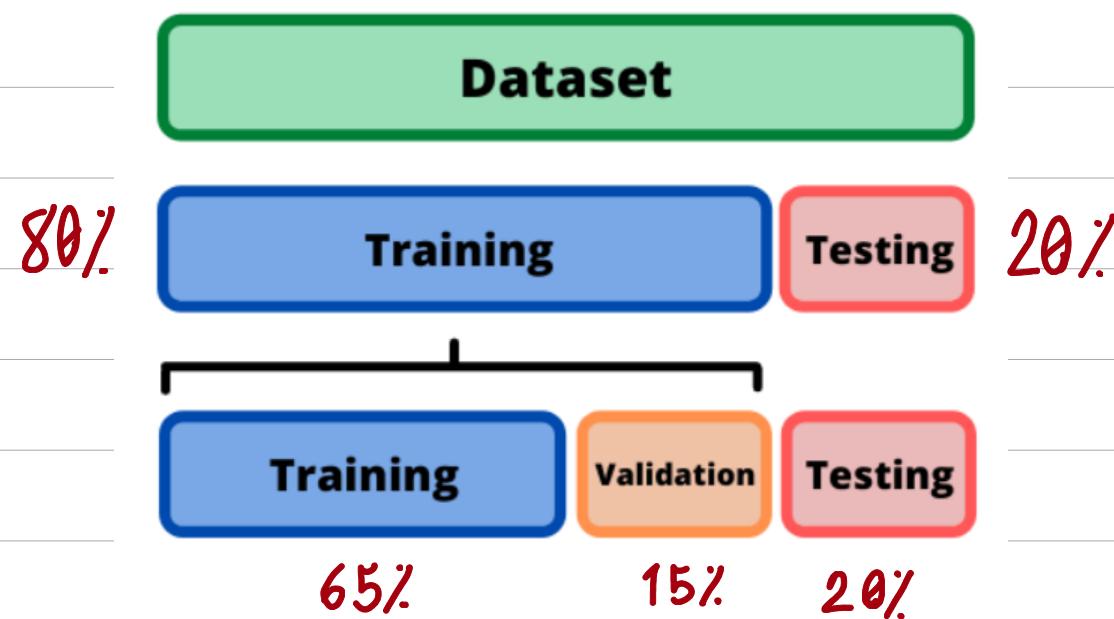
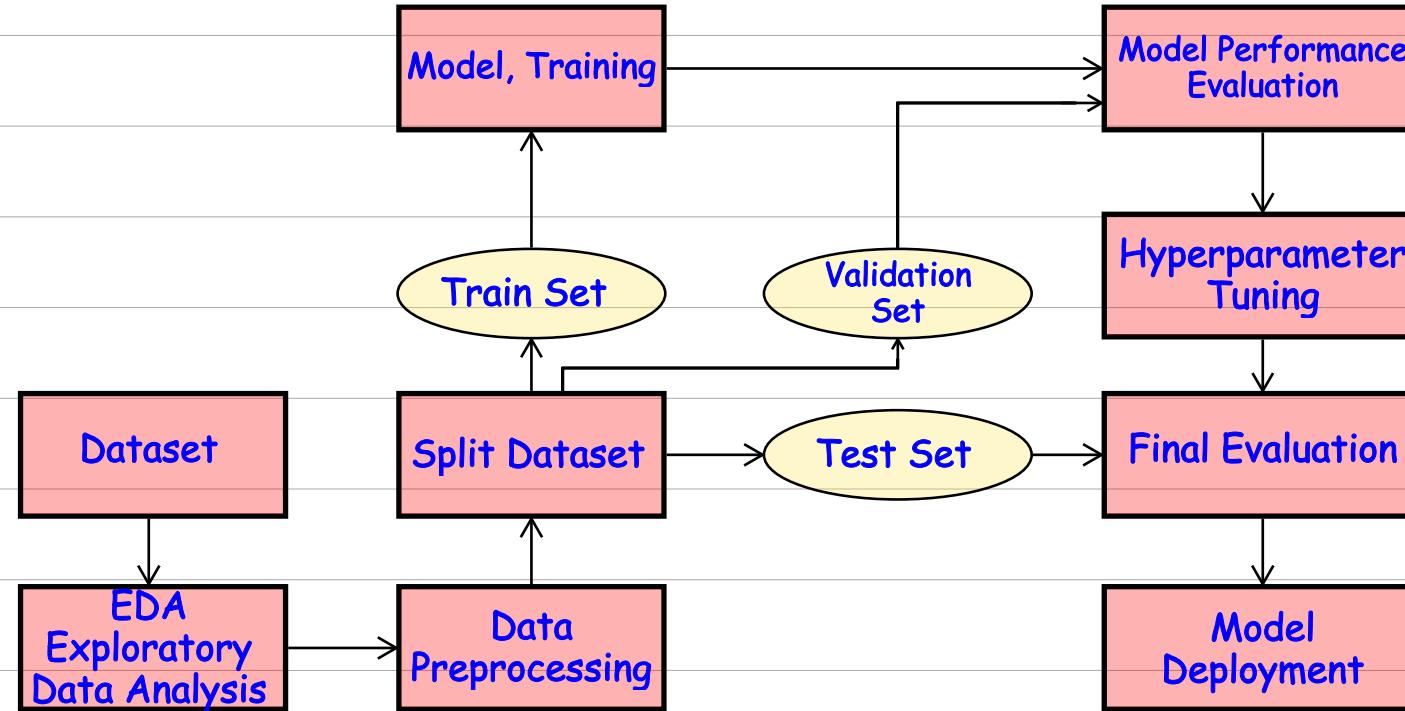
Continues پیوسته

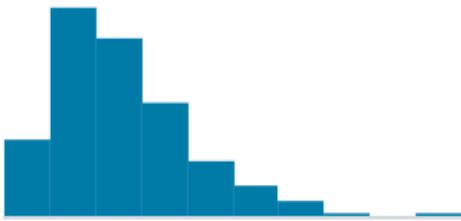
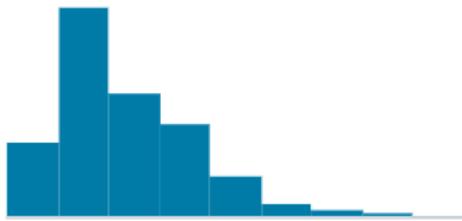
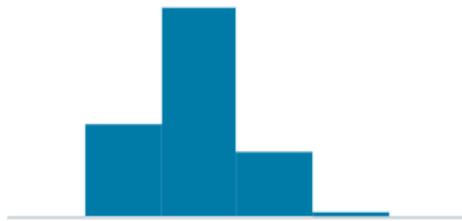
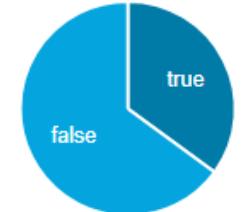
Discrete لسته

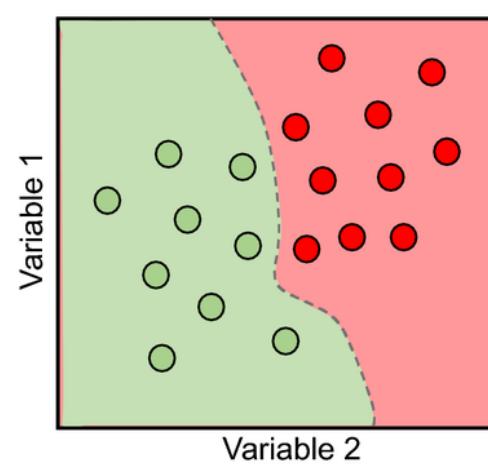
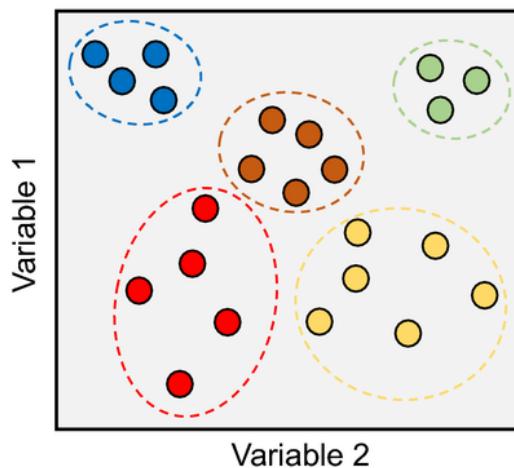
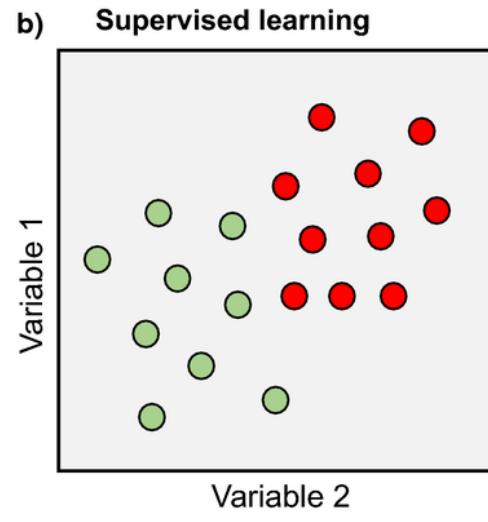
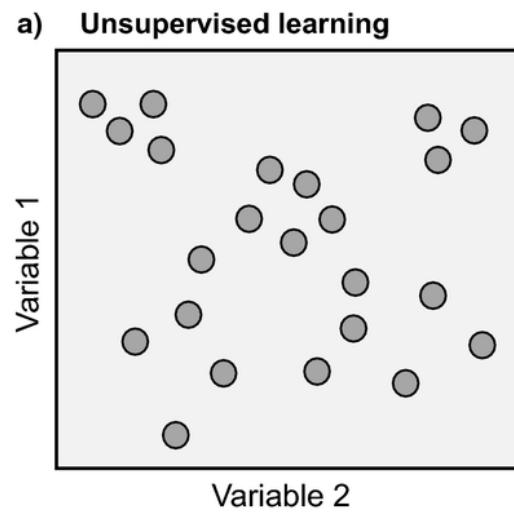
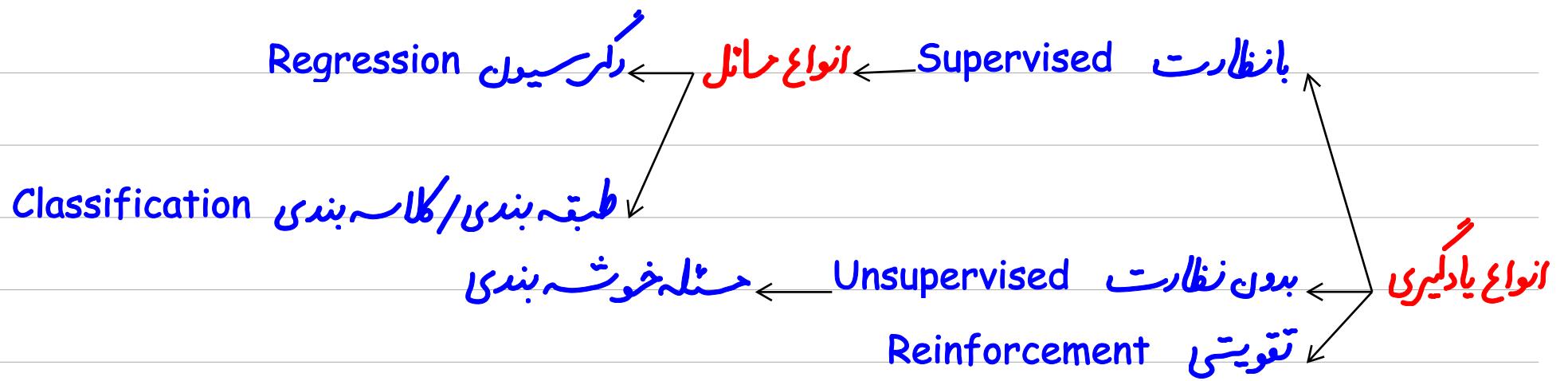
②

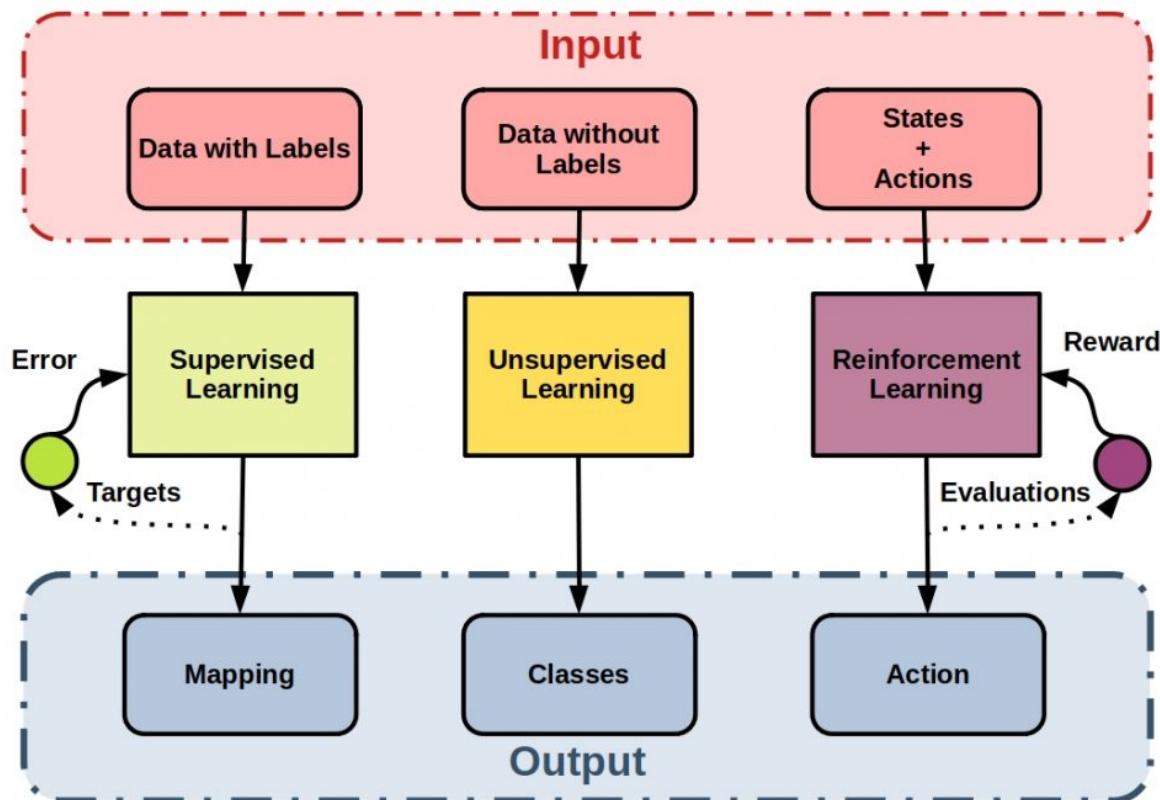
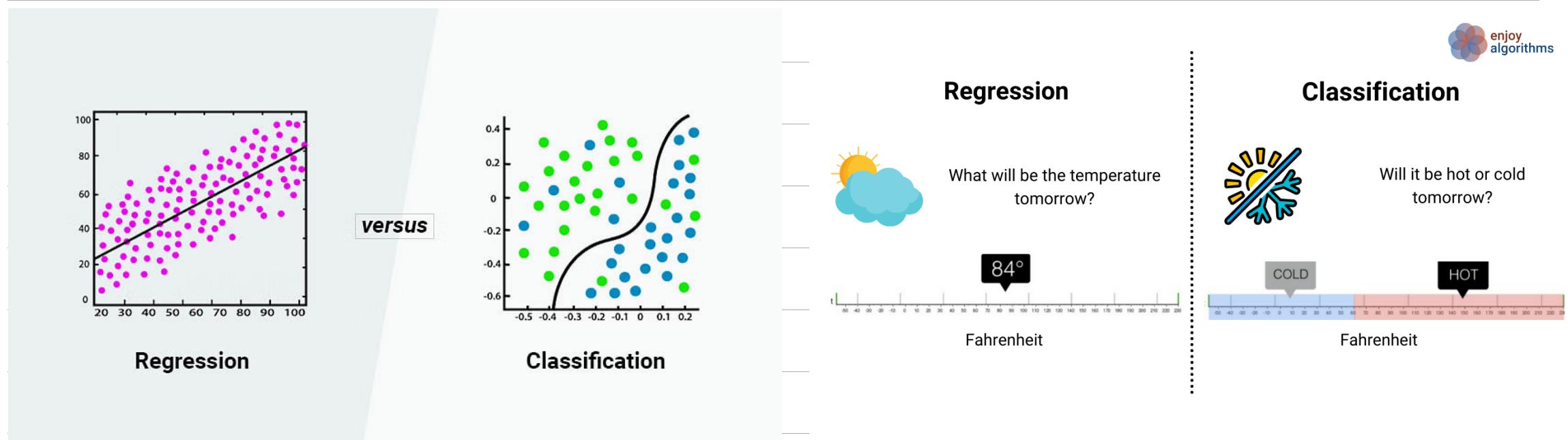
Numerical Data عددی

داده‌های کتگوریکال	داده‌های عددی	ویژگی
داده‌ها که به صورت دسته‌بندی‌های مشخص بیان می‌شوند	داده‌ها که به صورت اعداد بیان می‌شوند	تعریف
نامینال یا ترتیبی	پیوسته یا گسسته	نوع
جنسیت، رنگ، رتبه‌بندی، نوع بیمه	سن، قد، وزن، دما	مثال‌ها
غیرمعنادار	معنادار (جمع، تفریق، میانگین)	عملیات ریاضی
تجزیه و تحلیل فراوانی، تست‌های آماری کتگوریکال	تجزیه و تحلیل‌های کمی، رگرسیون	استفاده در تجزیه و تحلیل
نمودارهای میله‌ای، دایره‌ای، پله‌ای	نمودارهای میله‌ای، خطی، پراکندگی	تجسم

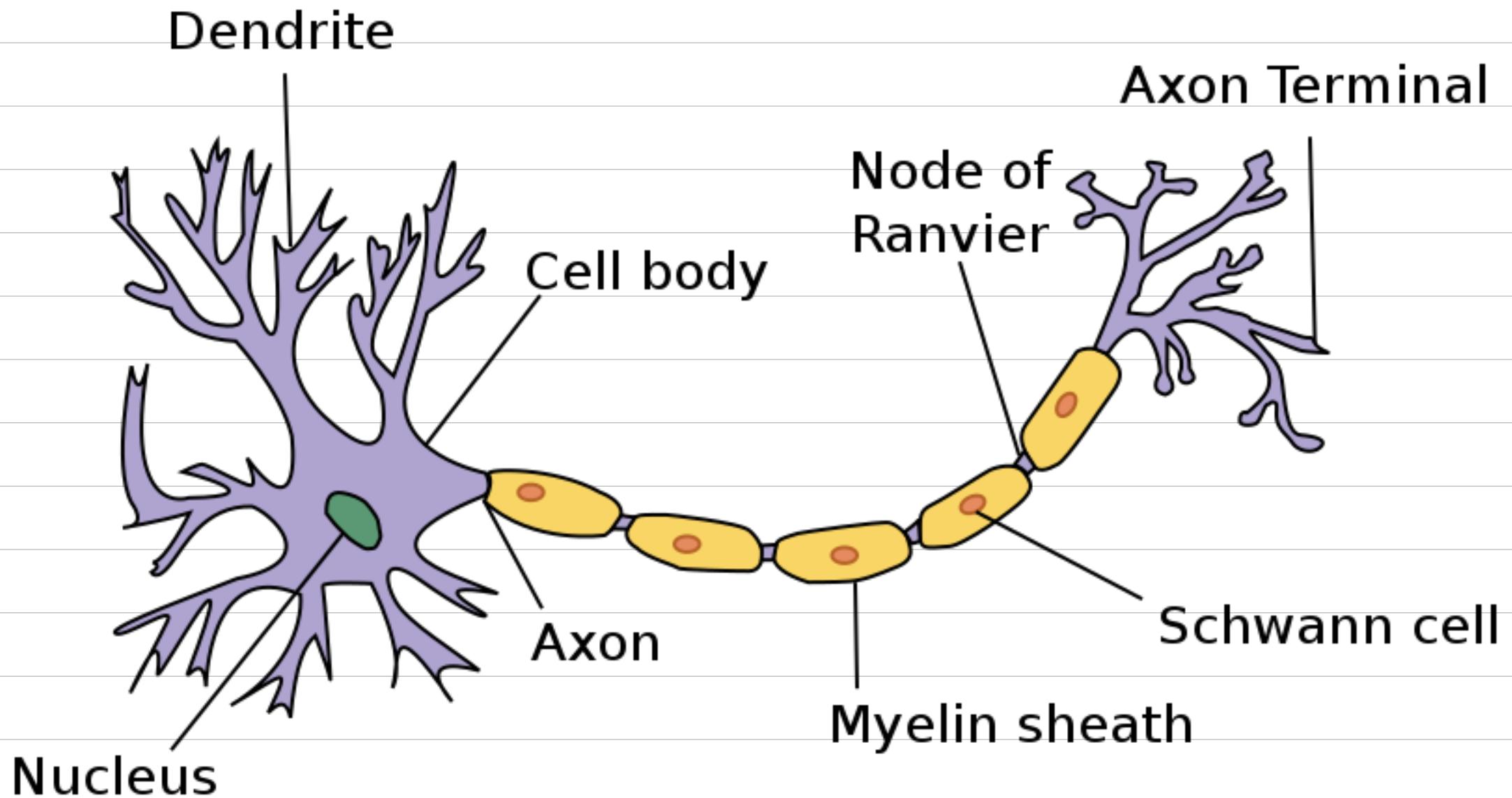


# price	# area	# bedrooms	✓ basement
Price of the Houses	Area of a House	Number of House Bedrooms	Weather has a basement
			
1.75m	1650	1	true 191 35%
13300000	7420	4	false 354 65%
12250000	8960	4	no
12250000	9960	3	no
12215000	7500	4	yes
11410000	7420	4	yes
10850000	7500	3	yes
10150000	8580	4	no
10150000	16200	5	no
9870000	8100	4	yes
9800000	5750	3	no

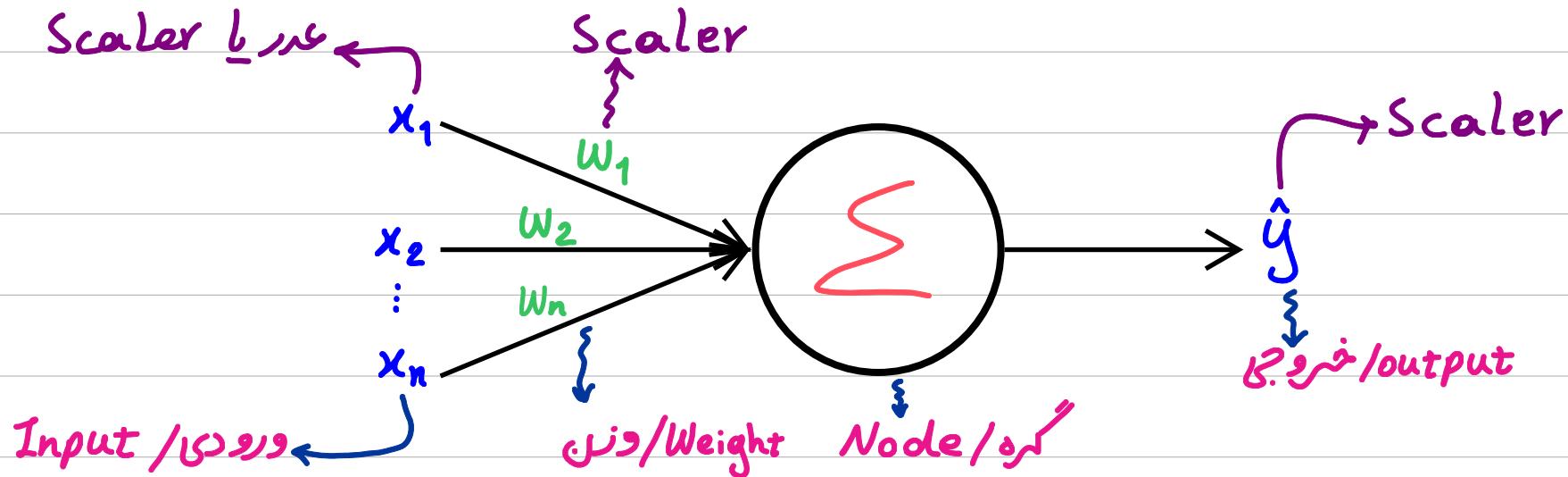




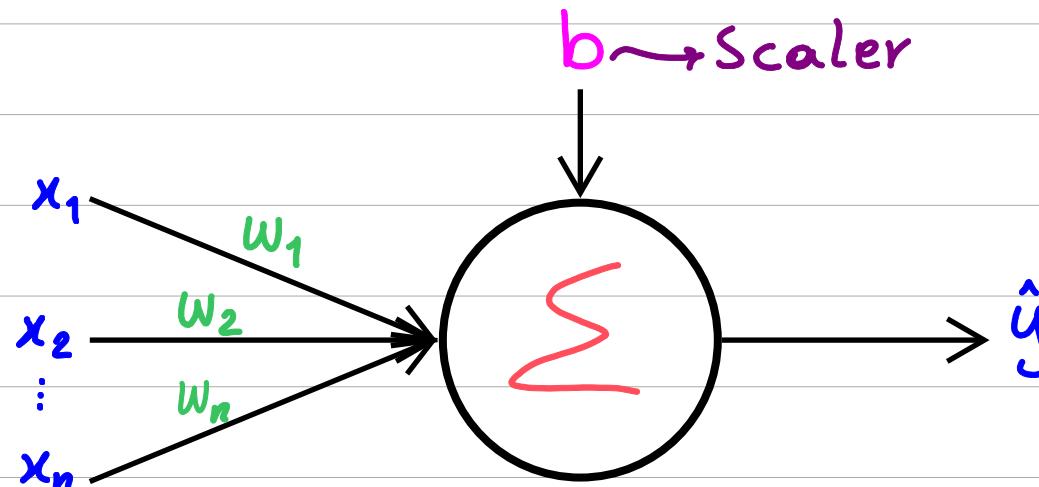
Neuron



Artificial Neuron

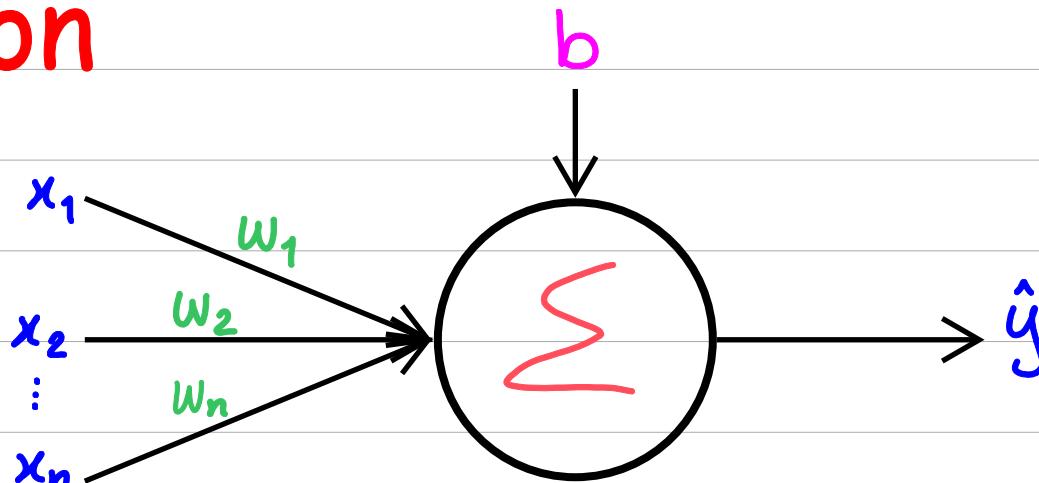


$$\hat{y} = x_1 \cdot w_1 + x_2 \cdot w_2 + \dots + x_n \cdot w_n = \sum_{i=1}^n x_i \cdot w_i$$



$$\hat{y} = x_1 \cdot w_1 + x_2 \cdot w_2 + \dots + x_n \cdot w_n + b = \sum_{i=1}^n x_i \cdot w_i + b$$

Vectorization



$$\hat{y} = x_1 \cdot w_1 + x_2 \cdot w_2 + \dots + x_n \cdot w_n + b = \sum_{i=1}^n x_i \cdot w_i + b$$

$$W \quad X \quad W^T \quad X$$

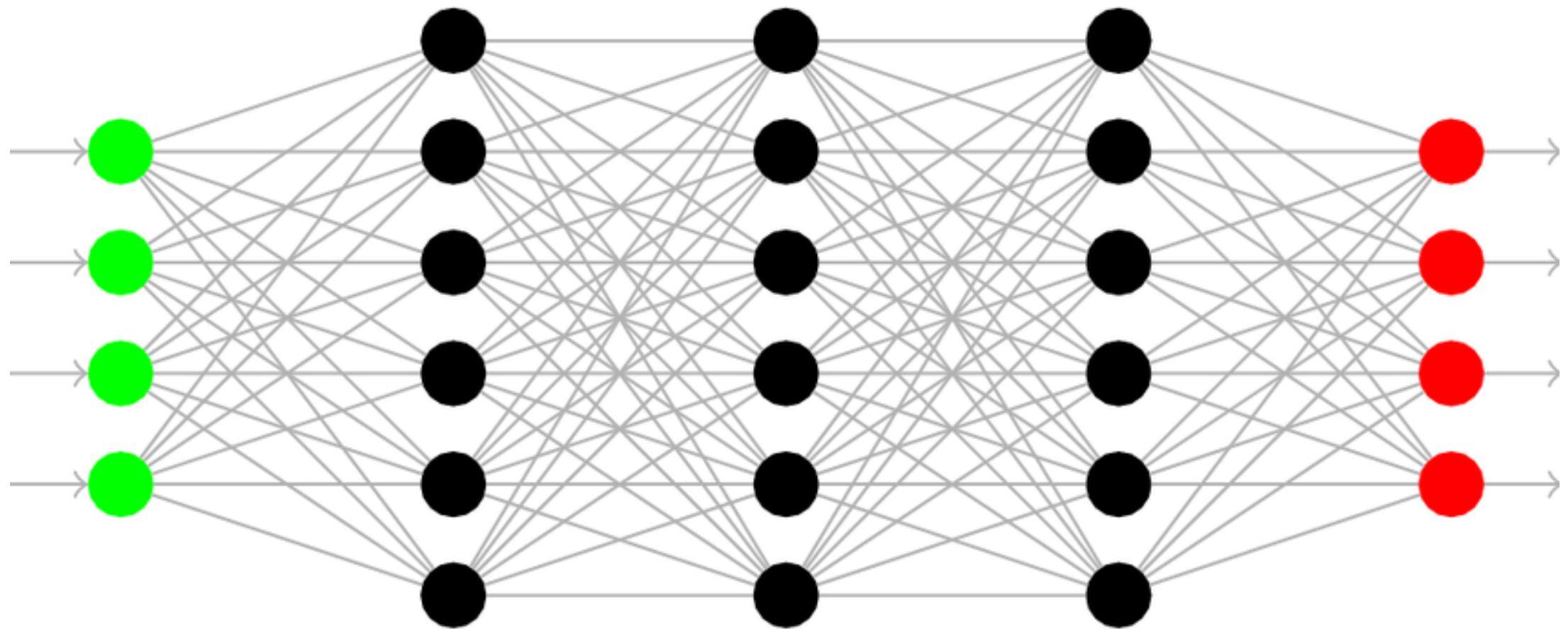
$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}_{n \times 1} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1} + b = \hat{y} \rightsquigarrow \begin{bmatrix} w_1 & w_2 & \dots & w_n \end{bmatrix}_{1 \times n} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1} + b = \hat{y}$$

:

$$\begin{bmatrix} a_{00} & a_{01} & \dots & a_{0n} \\ a_{10} & a_{11} & \dots & a_{1n} \\ \vdots & & & \\ a_{m0} & a_{m1} & \dots & a_{mn} \end{bmatrix}_{m \times n} \begin{bmatrix} a_{00} & a_{01} & \dots & a_{0p} \\ a_{10} & a_{11} & \dots & a_{1p} \\ \vdots & & & \\ a_{n0} & a_{n1} & \dots & a_{np} \end{bmatrix}_{n \times p} = \begin{bmatrix} a_{00} & a_{01} & \dots & a_{0p} \\ a_{10} & a_{11} & \dots & a_{1p} \\ \vdots & & & \\ a_{m0} & a_{m1} & \dots & a_{mp} \end{bmatrix}_{m \times p}$$

$=$

Neural Network Layers



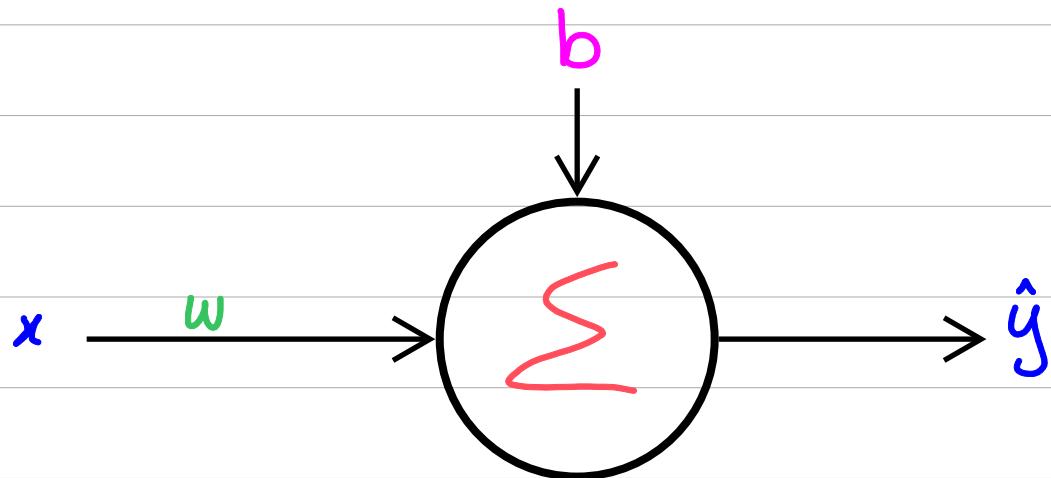
input node

hidden node

output node

If #Hidden layers $\geq 2 \rightarrow$ Deep Neural Network

Feed-Forward Neural Network (FFNN)



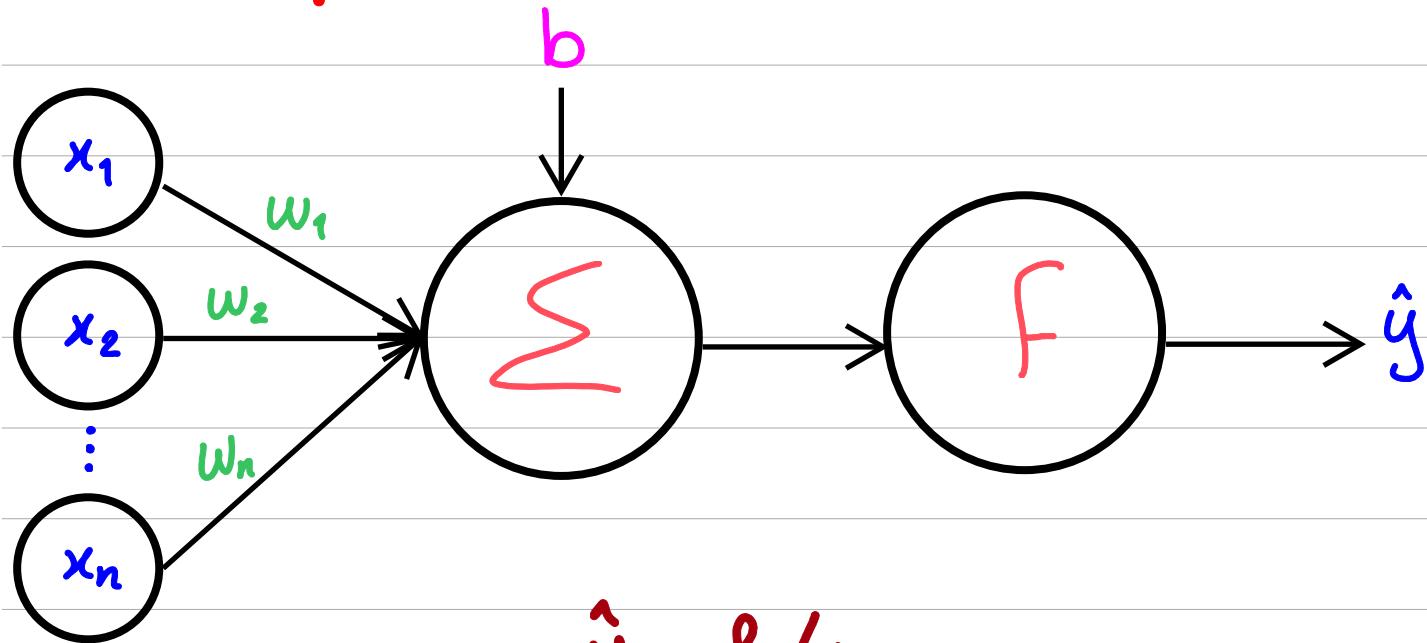
Size (m ²)	Price (k\$)	\hat{y}
800	200	208
1000	240	220
1200	280	232

$$w = 60 \quad b = 160000$$

$$\textcircled{1} \quad \hat{y} = wx = (60 \times 800) = 48,000 \quad 48,000 + \overbrace{160,000}^b = 208,000$$

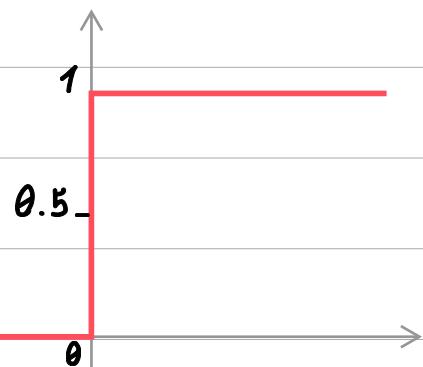
$$\textcircled{2} \quad \hat{y} = wx + b = (60 \times 1000) + 160,000 = 220,000 \quad \textcircled{3} \quad \hat{y} = wx + b = (1200 \times 60) + 160,000 = 232,000$$

Perceptron



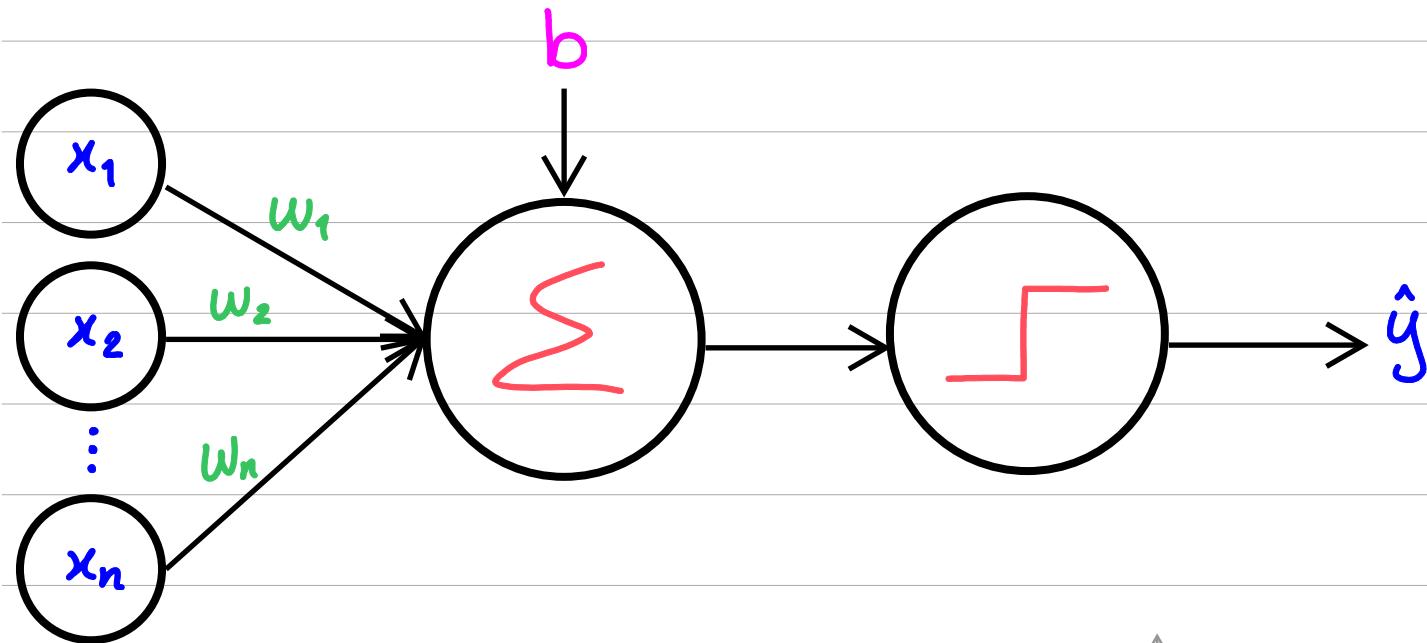
$$\hat{y} = f(w_1x_1 + w_2x_2 + \dots + w_nx_n + b) = f\left(\sum_{i=1}^n w_i x_i + b\right)$$

📌 Step Function:



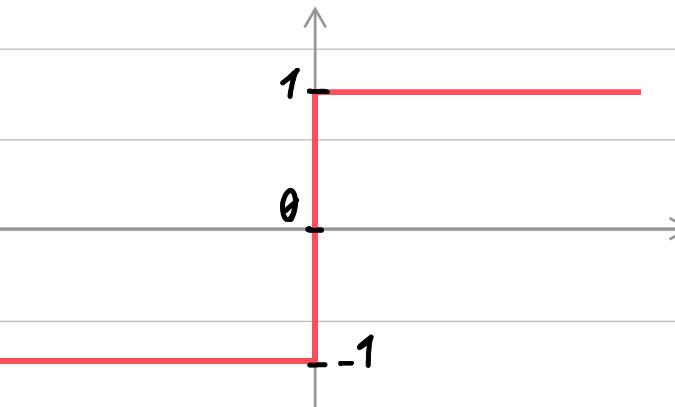
$$y = \begin{cases} 1 & x > \theta \\ 0 & x < \theta \end{cases}$$

Perceptron



📌 Sign Function:

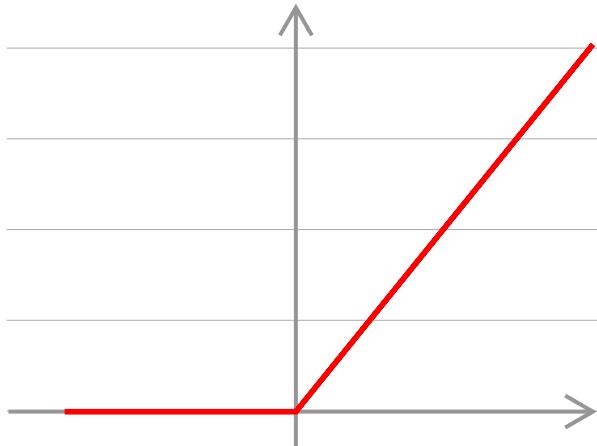
Heaviside



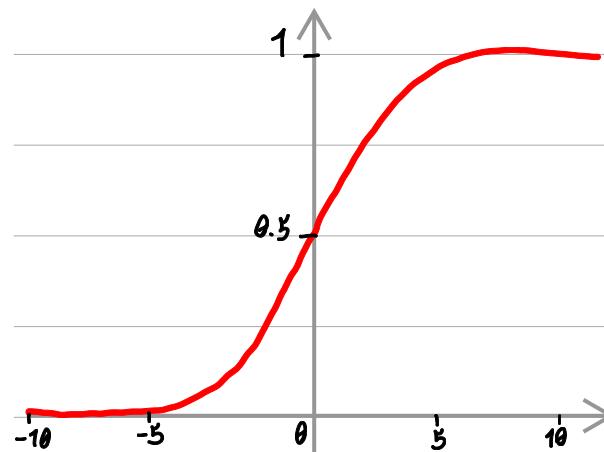
$$y = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$$

$$\hat{y} = \text{Step}(w_1x_1 + w_2x_2 + \dots + w_nx_n + b) = \text{Step}\left(\sum_{i=1}^n w_i x_i + b\right)$$

Activation Function

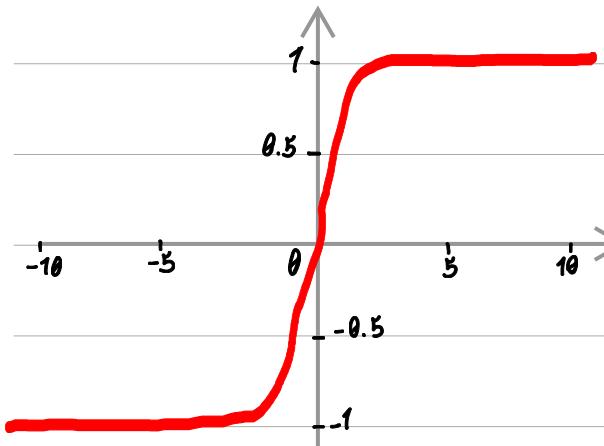


$$\text{ReLU}(z) = \max(0, z)$$



Sigmoid $\rightsquigarrow \sigma(z) = \frac{1}{1 + e^{-z}}$

Softmax $\rightsquigarrow \sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}}$



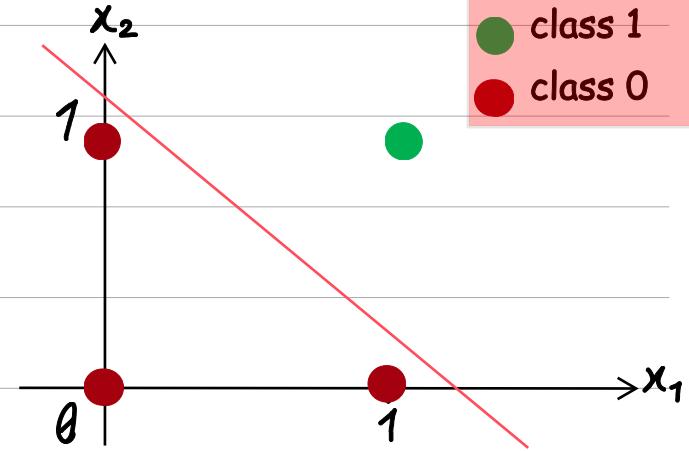
tanh $\rightsquigarrow \sigma(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$

Logic Gates

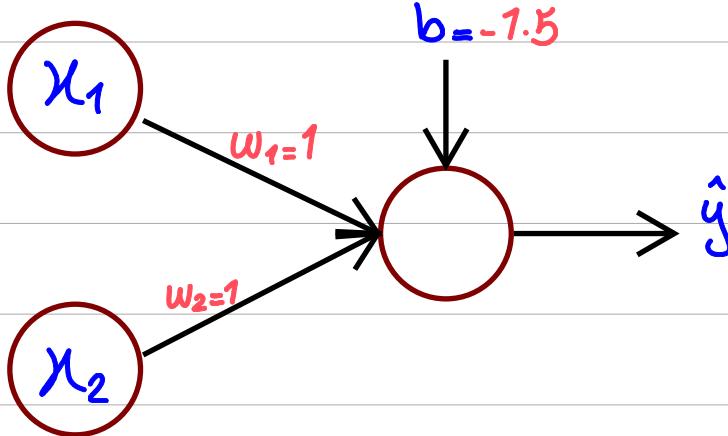
1. And &



$$\text{Step} = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases}$$



x_1	x_2	x_1 and x_2
T 1	T 1	T 1
F 0	T 1	F 0
T 1	F 0	F 0
F 0	F 0	F 0



$$①: \hat{y} = \text{Step}(\sum w_i x_i + b) = \text{Sign}((1 \times 1) + (1 \times 1) + (-1.5)) = 1$$

$$②: \hat{y} = \text{Step}((1 \times 0) + (1 \times 1) + (-1.5)) = 0$$

$$③: \hat{y} = \text{Step}((1 \times 1) + (0 \times 1) + (-1.5)) = 0$$

$$④: \hat{y} = \text{Step}((0 \times 1) + (0 \times 1) + (-1.5)) = 0$$

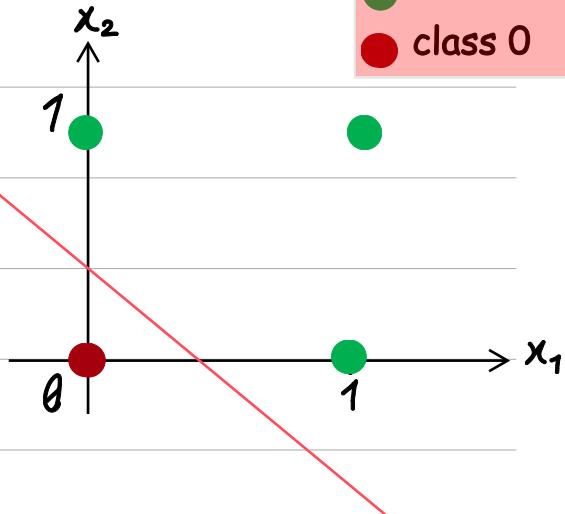
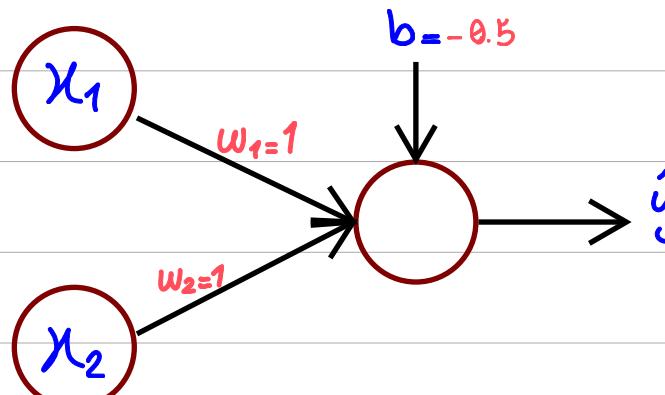
Logic Gates

class 1
class 0

2. Or |



x_1	x_2	$x_1 \text{ or } x_2$
T 1	T 1	T 1
F 0	T 1	T 1
T 1	F 0	T 1
F 0	F 0	F 0



$$\textcircled{1}: \hat{y} = \text{Step} \left(\cancel{(1 \times 1)} + \cancel{(1 \times 1)} + (-0.5) \right) = 1$$

$$\textcircled{2}: \hat{y} = \text{Step} \left(\cancel{(0 \times 1)} + \cancel{(1 \times 1)} + (-0.5) \right) = 1$$

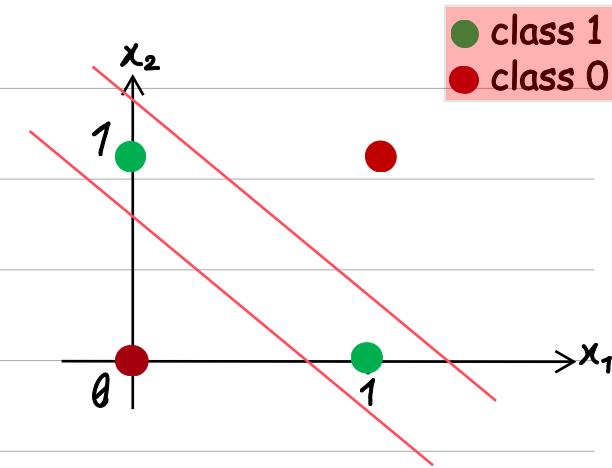
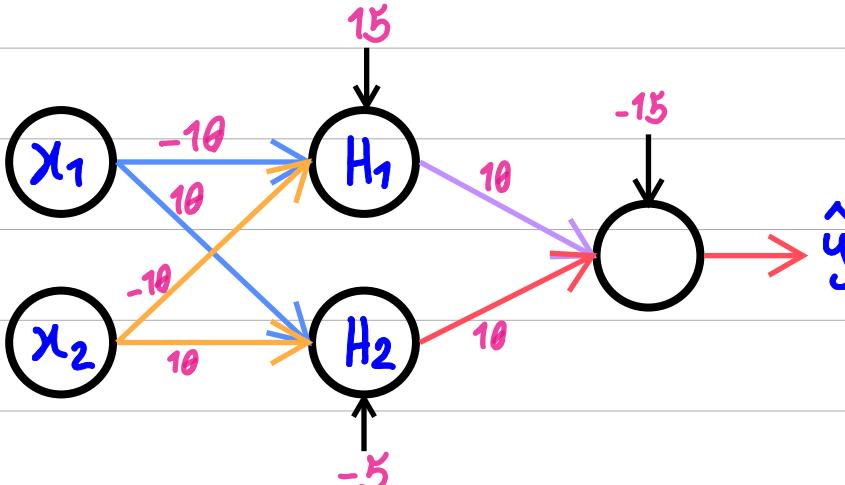
$$\textcircled{3}: \hat{y} = \text{Step} \left(\cancel{(1 \times 1)} + \cancel{(0 \times 1)} + (-0.5) \right) = 1$$

$$\textcircled{4}: \hat{y} = \text{Step} \left(\cancel{(0 \times 1)} + \cancel{(0 \times 1)} + (-0.5) \right) = 0$$

Multilayer Perceptron (MLP)

3. XOR

x_1	x_2	$x_1 \text{ xor } x_2$
T 1	T 1	F 0
F 0	T 1	T 1
T 1	F 0	T 1
F 0	F 0	F 0



$$\textcircled{1} \cdot H_1 = \text{Sigmoid} (\sum Wx + b) = \text{Sigmoid} ((-10 \times 1) + (-10 \times 1) + 15) = \text{Sigmoid} (-5) = 0.006$$

$$H_2 = \text{Sigmoid} (\sum Wx + b) = \text{Sigmoid} ((10 \times 1) + (10 \times 1) - 5) = \text{Sigmoid} (15) = 0.999$$

$$\hat{y} = \text{Sigmoid} ((10 \times 0.006) + (10 \times 0.999) - 15) = \text{Sigmoid} (-4.88) = 0.0007 \approx 0 \rightarrow \hat{y} = 0$$

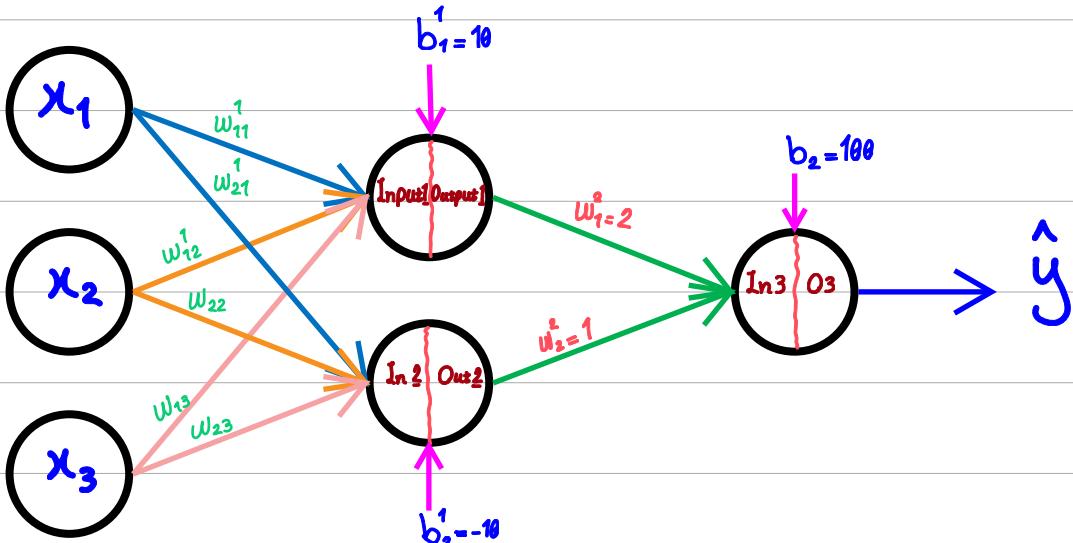
$$\textcircled{2} \cdot H_1 = \text{Sigmoid} ((-10 \times 0) + (-10 \times 1) + 15) = \text{Sigmoid} (5) = 0.993, H_2 = \text{Sigmoid} ((10 \times 0) + (10 \times 1) - 5) = \text{Sigmoid} (5) = 0.993$$

$$\hat{y} = \text{Sigmoid} ((10 \times 0.993) + (10 \times 0.993) - 15) = \text{Sigmoid} (4.86) = 0.992 \rightarrow \hat{y} = 1$$

$$\textcircled{3} \cdot \hat{y} = 1$$

$$\textcircled{4} \cdot \hat{y} = 0$$

Feed-Forward Neural Network (FFNN)



Size	#bedrooms	Age	Price
800	3	20	200

$w_1 = \begin{bmatrix} 0.01 & 50 & 5 \\ 0.02 & 25 & -25 \end{bmatrix}$, $b_1 = \begin{bmatrix} 10 \\ -10 \end{bmatrix}$,
 $w_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $b_2 = 100$

Hidden Layer
Activation Function = ReLU = max (0, number)

$$\text{Input 1} = \sum w_x + b = (0.01 \times 800) + (50 \times 3) + (5 \times 20) + 10 = 268 \xrightarrow{\text{AF}} \text{Output 1} = \text{ReLU}(268) = 268$$

$$\text{Input 2} = (0.02 \times 800) + (25 \times 3) + (-2.5 \times 20) + (-10) = 31 \xrightarrow{\text{AF}} \text{Output 2} = \text{ReLU}(31) = 31$$

$$\text{In 3} = (2 \times 268) + (1 \times 31) + 100 = 667 \rightsquigarrow \text{Output 3} = \hat{y} = 667 \quad \text{=} \quad \text{?}$$

Feed-Forward Neural Network (FFNN)



: Normalization \leadsto Value/max_value

$$\text{Size} = 800/1000 = 0.8$$

$$\text{Bedrooms} = 3/5 = 0.6$$

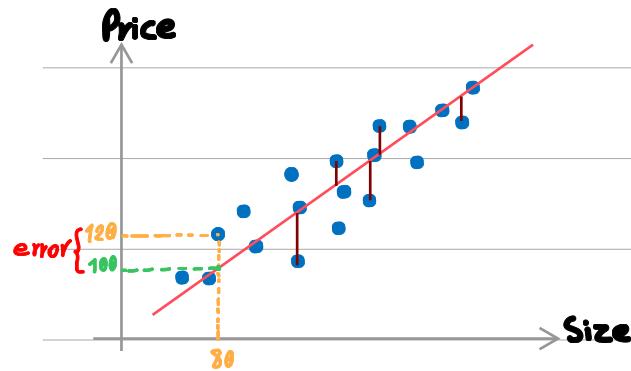
$$\text{Age} = 28/50 = 0.4$$

$$\text{In1} = (0.01 \times 0.8) + (50 \times 0.6) + (5 \times 0.4) + 10 = 42.008 \xrightarrow{\text{AF}} \text{Output1} = \text{ReLU}(42.008) = 42.008$$

$$\text{In2} = (0.02 \times 0.8) + (25 \times 0.6) + (-2.5 \times 0.4) + (-10) = 4.016 \xrightarrow{\text{AF}} \text{Output2} = \text{ReLU}(4.016) = 4.016$$

$$\text{In3} = (2 \times 42.008) + (1 \times 4.016) + 100 = 188.032 \xrightarrow{\text{AF}} \text{Output3} = \hat{y} = 188.032$$

Loss Function \rightsquigarrow Error \rightsquigarrow به اختلاف بین مقدار واقعی و پیش‌بینی مدل خطای کویند.



Regression Loss

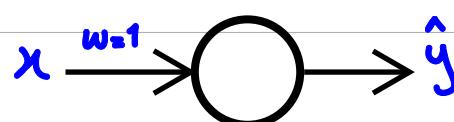
$$\begin{cases} 1. \text{MSE} = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i)^2 \\ 2. \text{MAE} = \frac{1}{m} \sum_{i=1}^m |\hat{y}_i - y_i| \\ 3. \text{RMSE} = \sqrt{\text{MSE}} \\ 4. \text{R2-Score} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \end{cases}$$

?

$\sum_{i=1}^n (y_i - \hat{y}_i)^2 \rightarrow SS_R$ (SUM of Squares for Regression)

$\sum_{i=1}^n (y_i - \bar{y})^2 \rightarrow SS_M$ (Sum of Squares for Model)

x	y	\hat{y}
1	2	1
2	4	2
3	6	3
4	8	4



$$MSE = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i)^2 = \frac{1}{4} [(1-2)^2 + (2-4)^2 + (3-6)^2 + (4-8)^2] = \frac{30}{4} = 7.5$$

$$R^2 = 1 - \frac{SS_R}{SS_M} = 1 - \frac{30}{20} = -0.5$$

؟ ☺ ۱۸ نویسخواه - ۰.۵

$$SS_R = (2-1)^2 + (4-2)^2 + (6-3)^2 + (8-4)^2 = 30$$

$$SS_M = (2-5)^2 + (4-5)^2 + (6-5)^2 + (8-5)^2 = 20$$

$$\bar{y} = \frac{2+4+6+8}{4} = 5$$

جیہے مارکی باسٹر مدل خوب ہے؟ R2

R2 = 1: Perfect fit

0 < R2 < 1: Good fit

R2 = 0: equivalent to using the mean

R2 < 0: Worse than using the mean

Thanks for
your attention