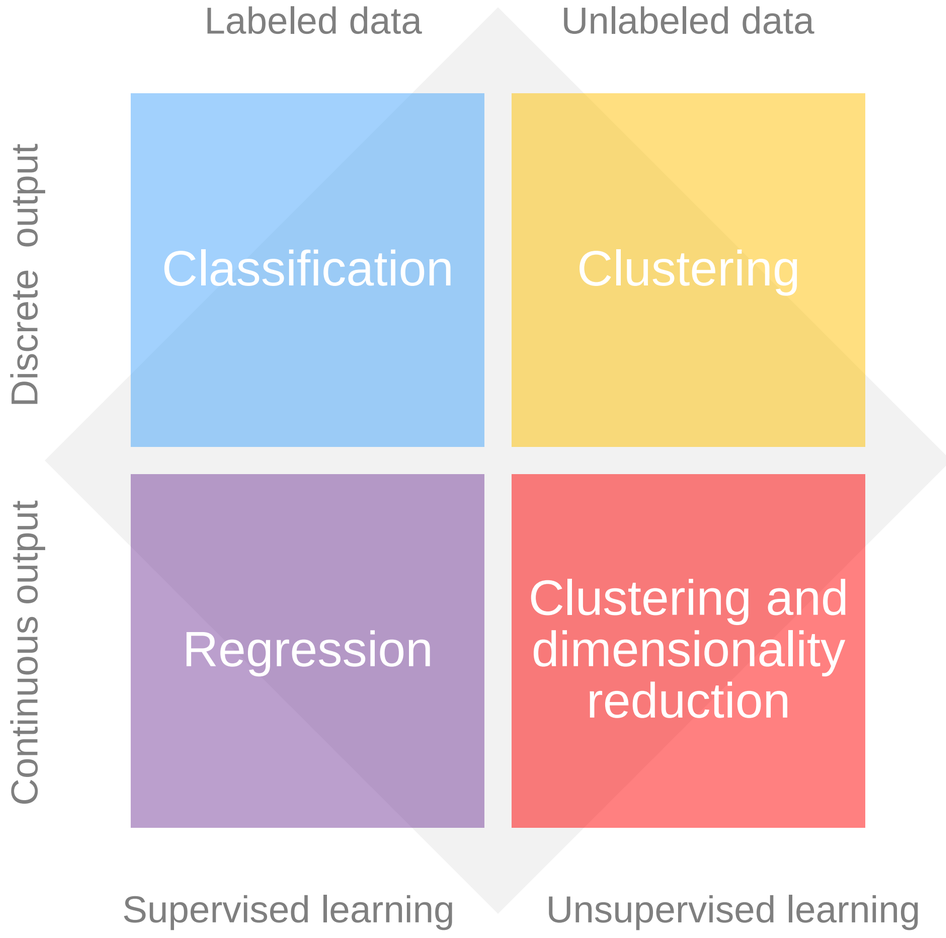
CREATIVE DATA MINING

Introduction to Supervised Machine Learning

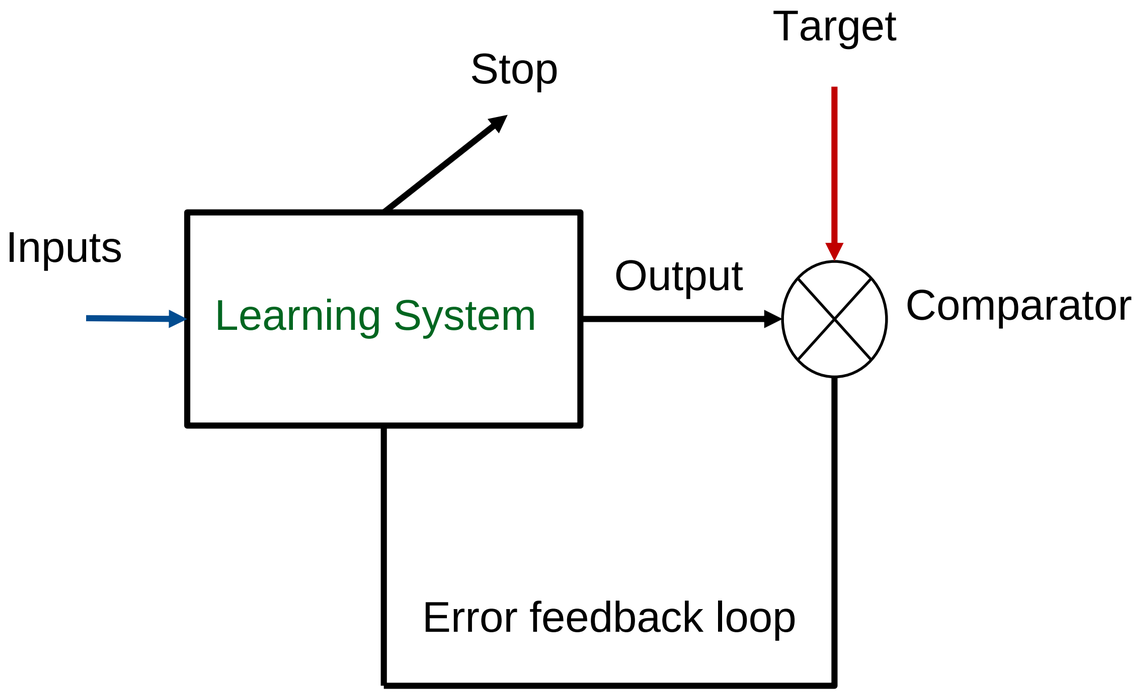
12.03.2018

Dr. Varun OJHA

Danielle GRIEGO



Supervised Learning Systems



Terms to remember:

Input

Target (Known output)

Output (System's output)

Feedback Loop (Training iteration)

Learning System (Model)

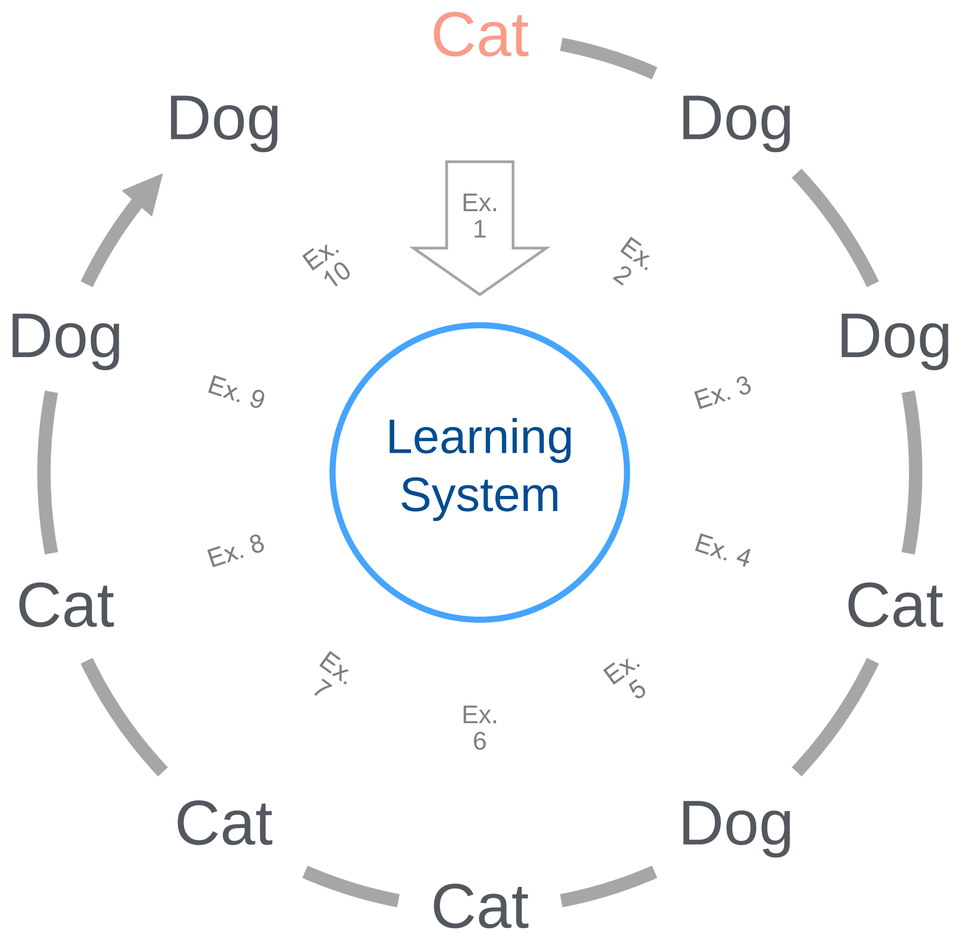
Stop (When to stop learning)

Data (labeled)

Discrete labeled data

#	Inputs		Target
	Length (cm)	Weight (kg)	Animal
Ex. 1	23.2	3.2	Cat
Ex. 2	70.9	19.5	Dog
Ex. 3	60.5	18.51	Dog
Ex. 4	24.5	4.6	Cat
Ex. 5	110.0	35.83	Dog
Ex. 6	23.8	3.7	Cat
Ex. 7	25.8	4.5	Cat
Ex. 8	24.7	4.9	Cat
Ex. 9	85.8	25.6	Dog
Ex. 10	78.8	20.33	Dog

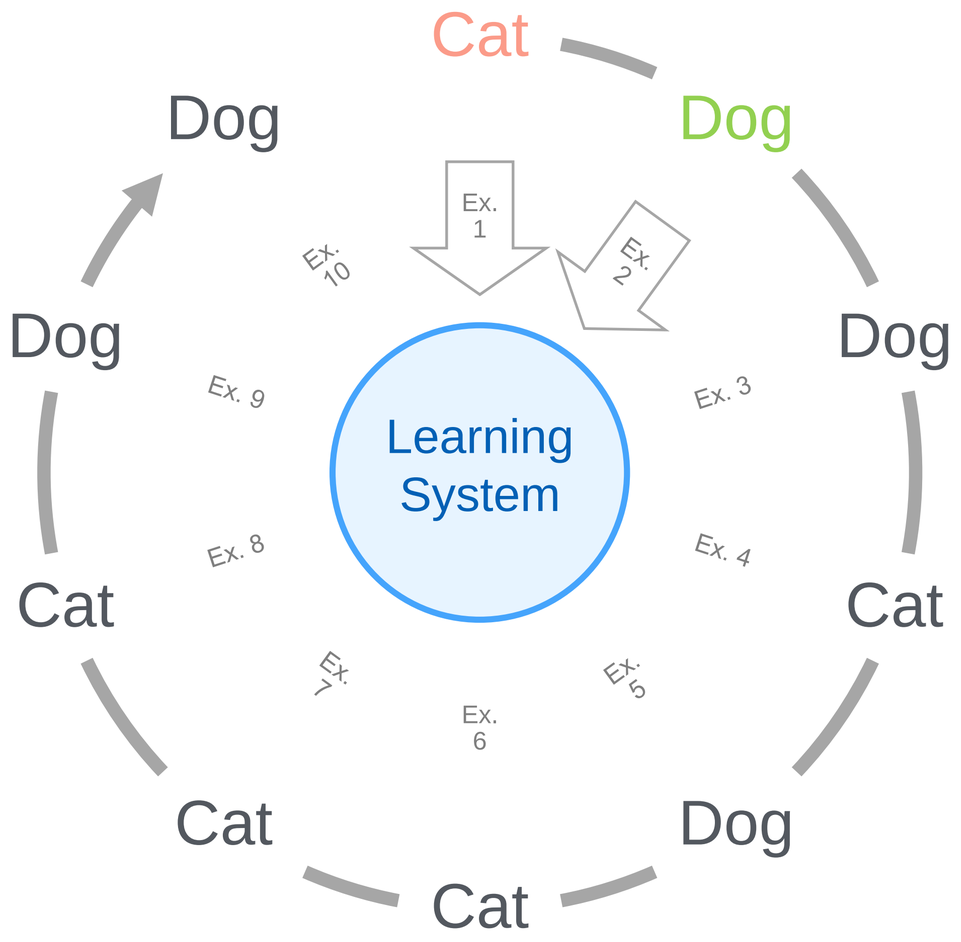
Data samples are labeled



Data (labeled)

Discrete labeled data

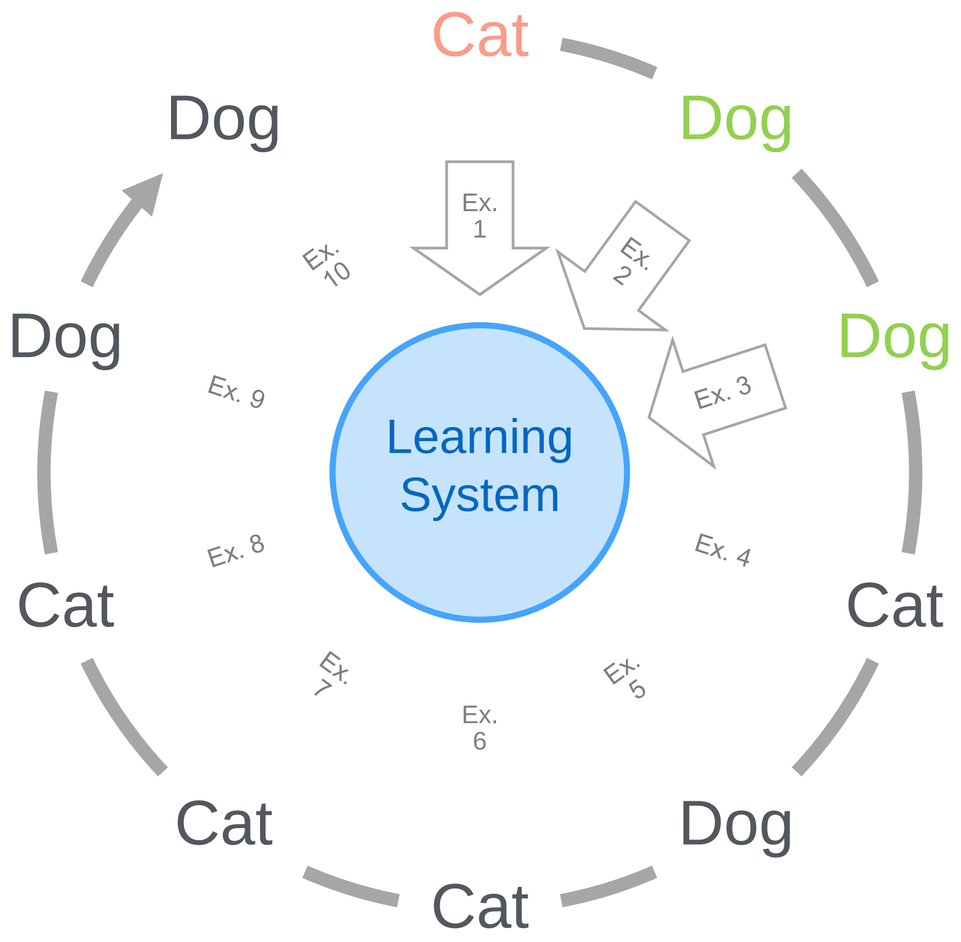
#	Inputs		Target
	Height (cm)	Weight (kg)	Animal
Ex. 1	23.2	3.2	Cat
Ex. 2	70.9	19.5	Dog
Ex. 3	60.5	18.51	Dog
Ex. 4	24.5	4.6	Cat
Ex. 5	110.0	35.83	Dog
Ex. 6	23.8	3.7	Cat
Ex. 7	25.8	4.5	Cat
Ex. 8	24.7	4.9	Cat
Ex. 9	85.8	25.6	Dog
Ex. 10	78.8	20.33	Dog



Data (labeled)

Discrete labeled data

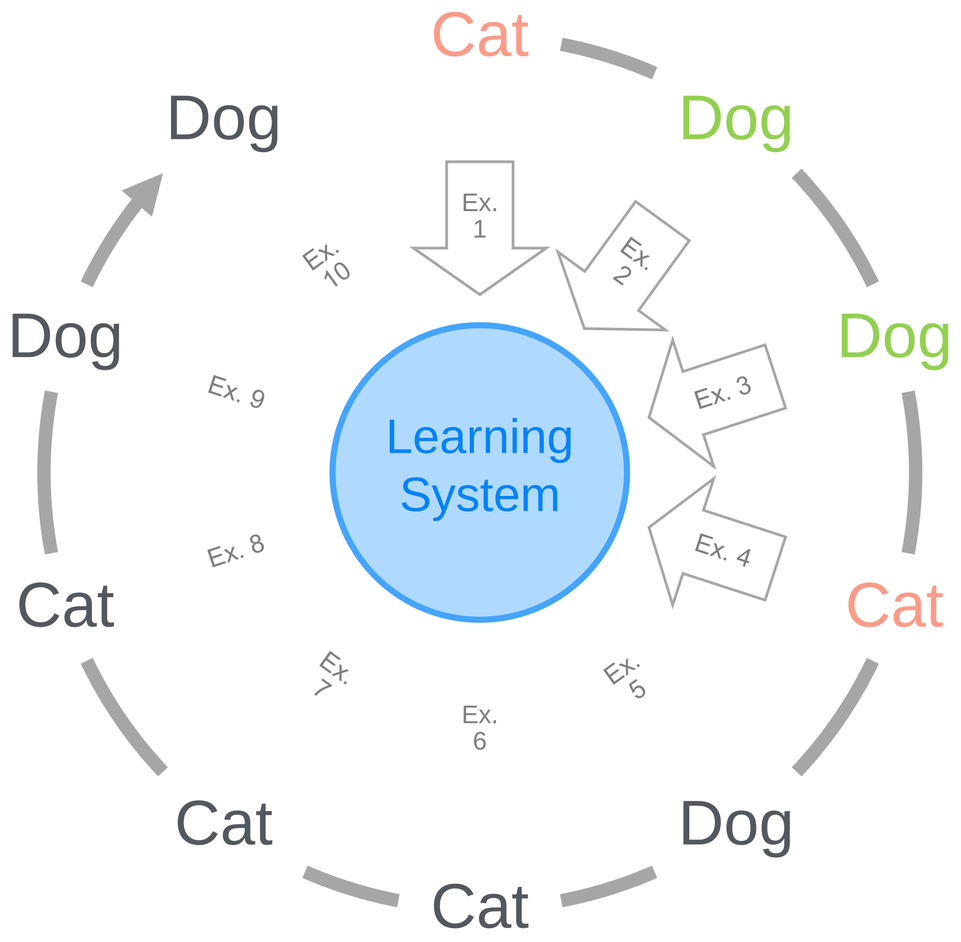
#	Inputs		Target
	Height (cm)	Weight (kg)	Animal
Ex. 1	23.2	3.2	Cat
Ex. 2	70.9	19.5	Dog
Ex. 3	60.5	18.51	Dog
Ex. 4	24.5	4.6	Cat
Ex. 5	110.0	35.83	Dog
Ex. 6	23.8	3.7	Cat
Ex. 7	25.8	4.5	Cat
Ex. 8	24.7	4.9	Cat
Ex. 9	85.8	25.6	Dog
Ex. 10	78.8	20.33	Dog



Data (labeled)

Discrete labeled data

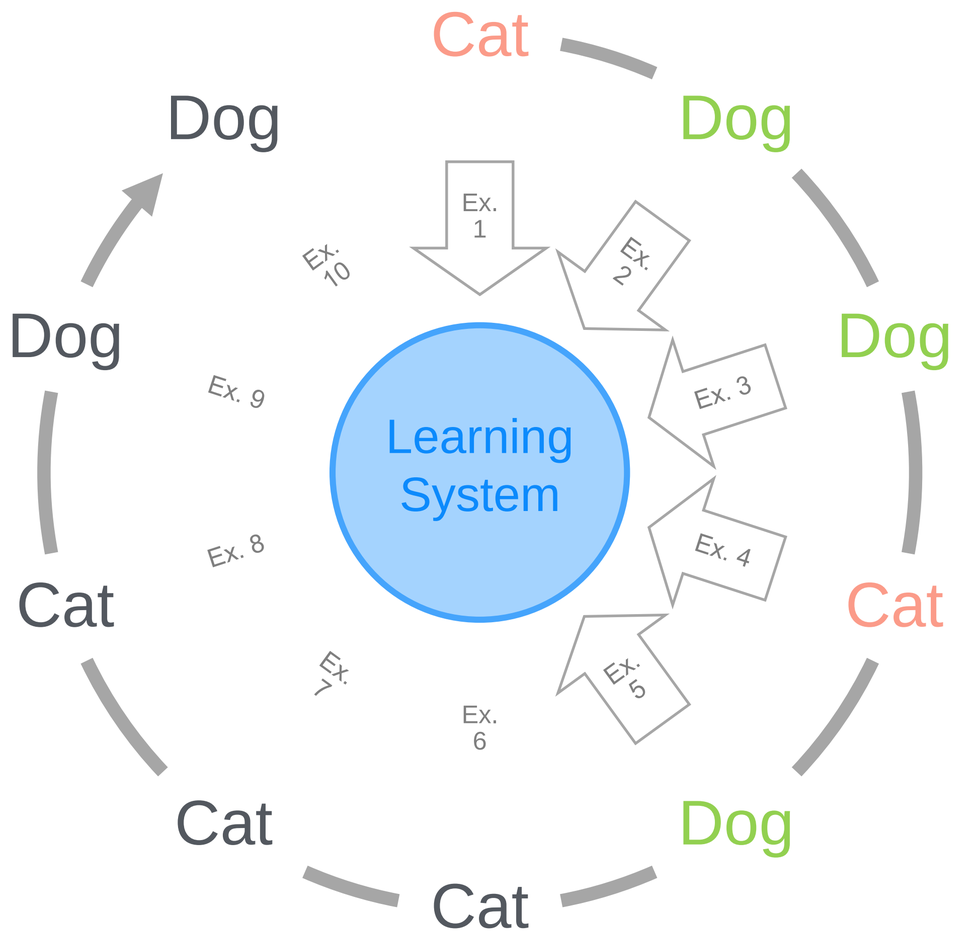
#	Inputs		Target
	Height (cm)	Weight (kg)	Animal
Ex. 1	23.2	3.2	Cat
Ex. 2	70.9	19.5	Dog
Ex. 3	60.5	18.51	Dog
Ex. 4	24.5	4.6	Cat
Ex. 5	110.0	35.83	Dog
Ex. 6	23.8	3.7	Cat
Ex. 7	25.8	4.5	Cat
Ex. 8	24.7	4.9	Cat
Ex. 9	85.8	25.6	Dog
Ex. 10	78.8	20.33	Dog



Data (labeled)

Discrete labeled data

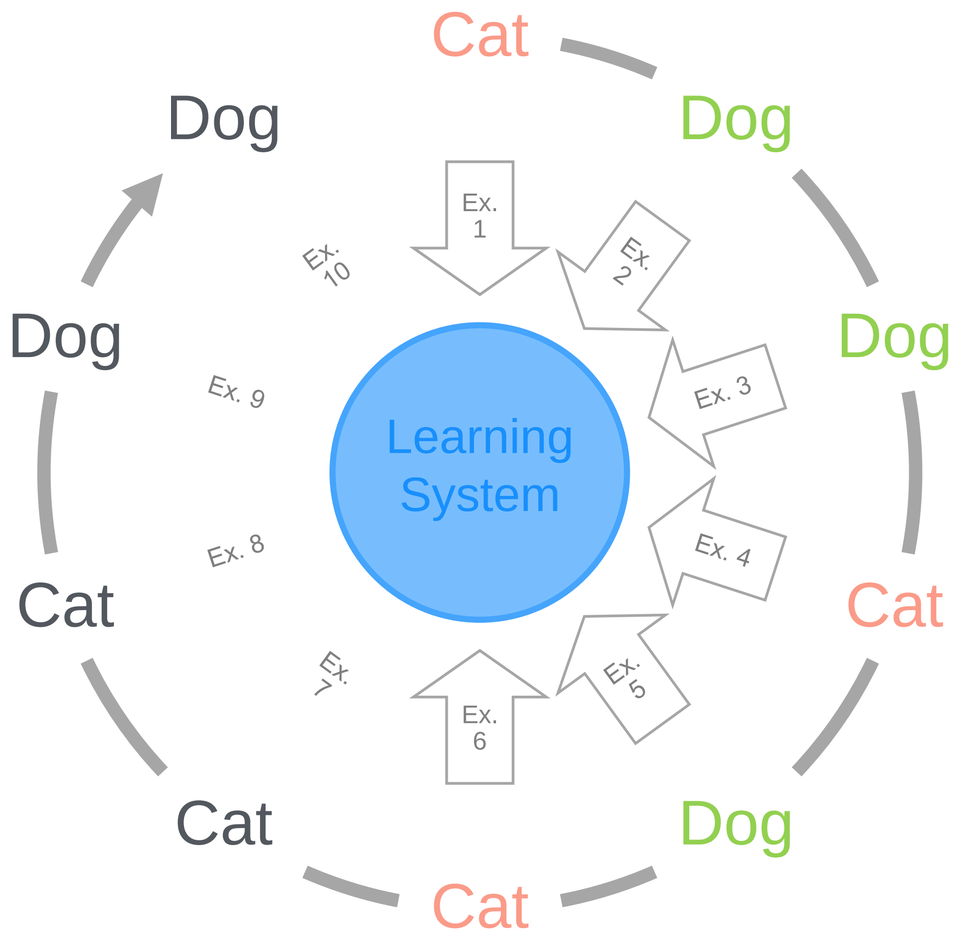
#	Inputs		Target
	Height (cm)	Weight (kg)	Animal
Ex. 1	23.2	3.2	Cat
Ex. 2	70.9	19.5	Dog
Ex. 3	60.5	18.51	Dog
Ex. 4	24.5	4.6	Cat
Ex. 5	110.0	35.83	Dog
Ex. 6	23.8	3.7	Cat
Ex. 7	25.8	4.5	Cat
Ex. 8	24.7	4.9	Cat
Ex. 9	85.8	25.6	Dog
Ex. 10	78.8	20.33	Dog



Data (labeled)

Discrete labeled data

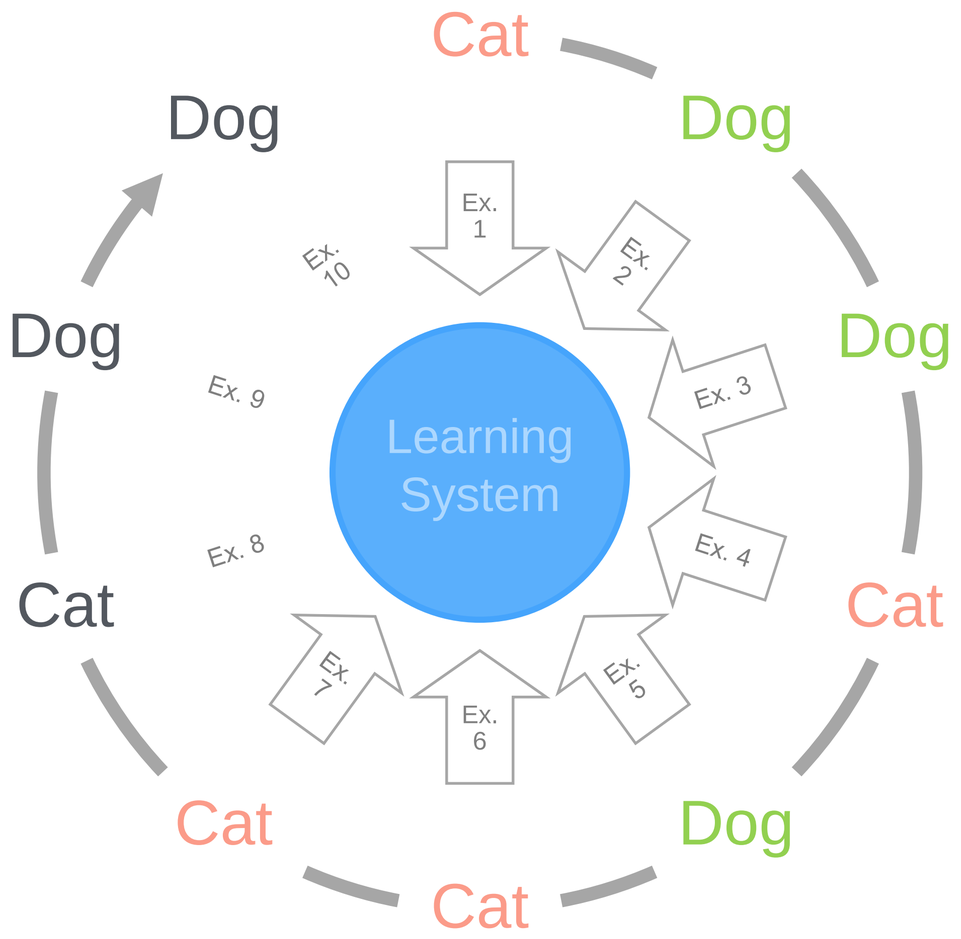
#	Inputs		Target
	Height (cm)	Weight (kg)	Animal
Ex. 1	23.2	3.2	Cat
Ex. 2	70.9	19.5	Dog
Ex. 3	60.5	18.51	Dog
Ex. 4	24.5	4.6	Cat
Ex. 5	110.0	35.83	Dog
Ex. 6	23.8	3.7	Cat
Ex. 7	25.8	4.5	Cat
Ex. 8	24.7	4.9	Cat
Ex. 9	85.8	25.6	Dog
Ex. 10	78.8	20.33	Dog



Data (labeled)

Discrete labeled data

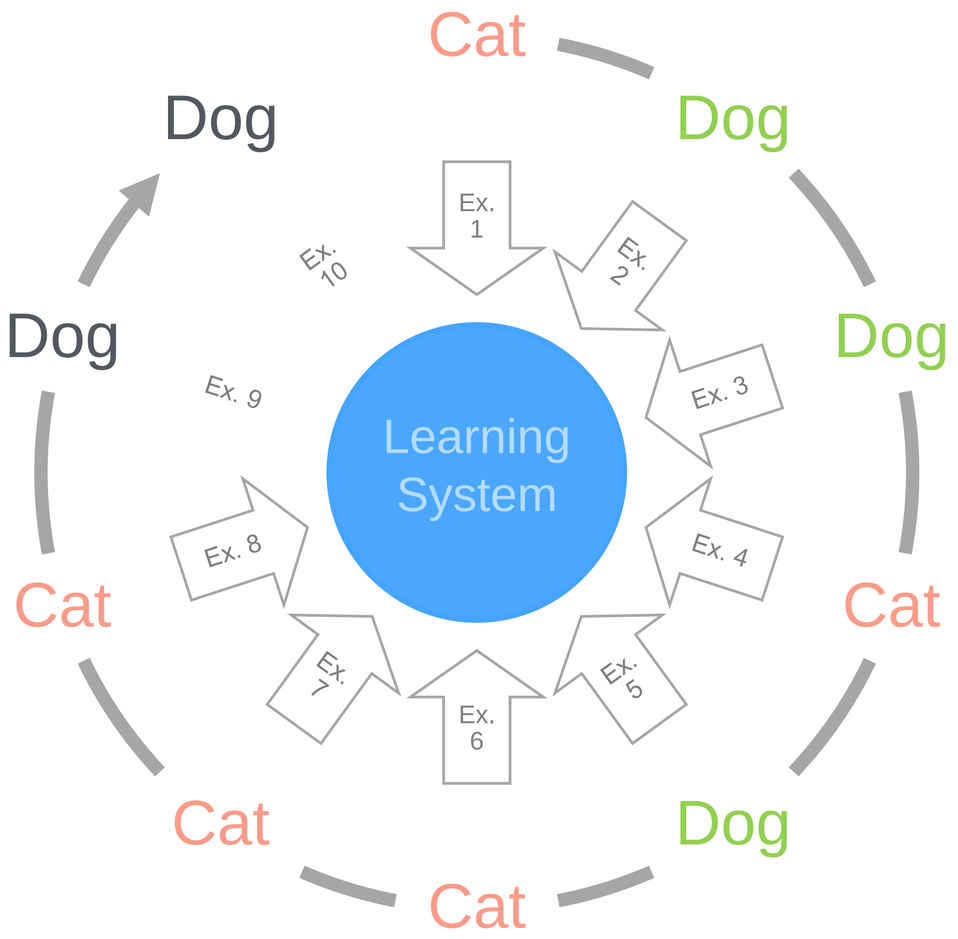
#	Inputs		Target
	Height (cm)	Weight (kg)	Animal
Ex. 1	23.2	3.2	Cat
Ex. 2	70.9	19.5	Dog
Ex. 3	60.5	18.51	Dog
Ex. 4	24.5	4.6	Cat
Ex. 5	110.0	35.83	Dog
Ex. 6	23.8	3.7	Cat
Ex. 7	25.8	4.5	Cat
Ex. 8	24.7	4.9	Cat
Ex. 9	85.8	25.6	Dog
Ex. 10	78.8	20.33	Dog



Data (labeled)

Discrete labeled data

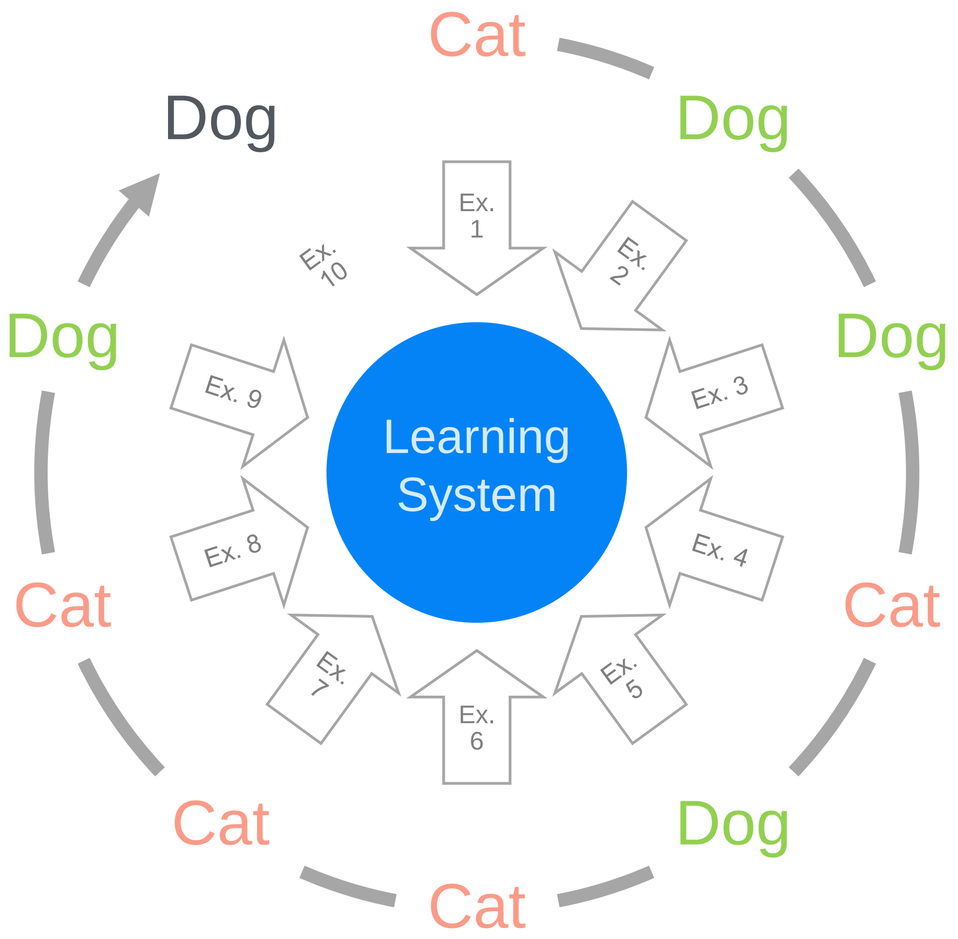
#	Inputs		Target
	Height (cm)	Weight (kg)	Animal
Ex. 1	23.2	3.2	Cat
Ex. 2	70.9	19.5	Dog
Ex. 3	60.5	18.51	Dog
Ex. 4	24.5	4.6	Cat
Ex. 5	110.0	35.83	Dog
Ex. 6	23.8	3.7	Cat
Ex. 7	25.8	4.5	Cat
Ex. 8	24.7	4.9	Cat
Ex. 9	85.8	25.6	Dog
Ex. 10	78.8	20.33	Dog



Data (labeled)

Discrete labeled data

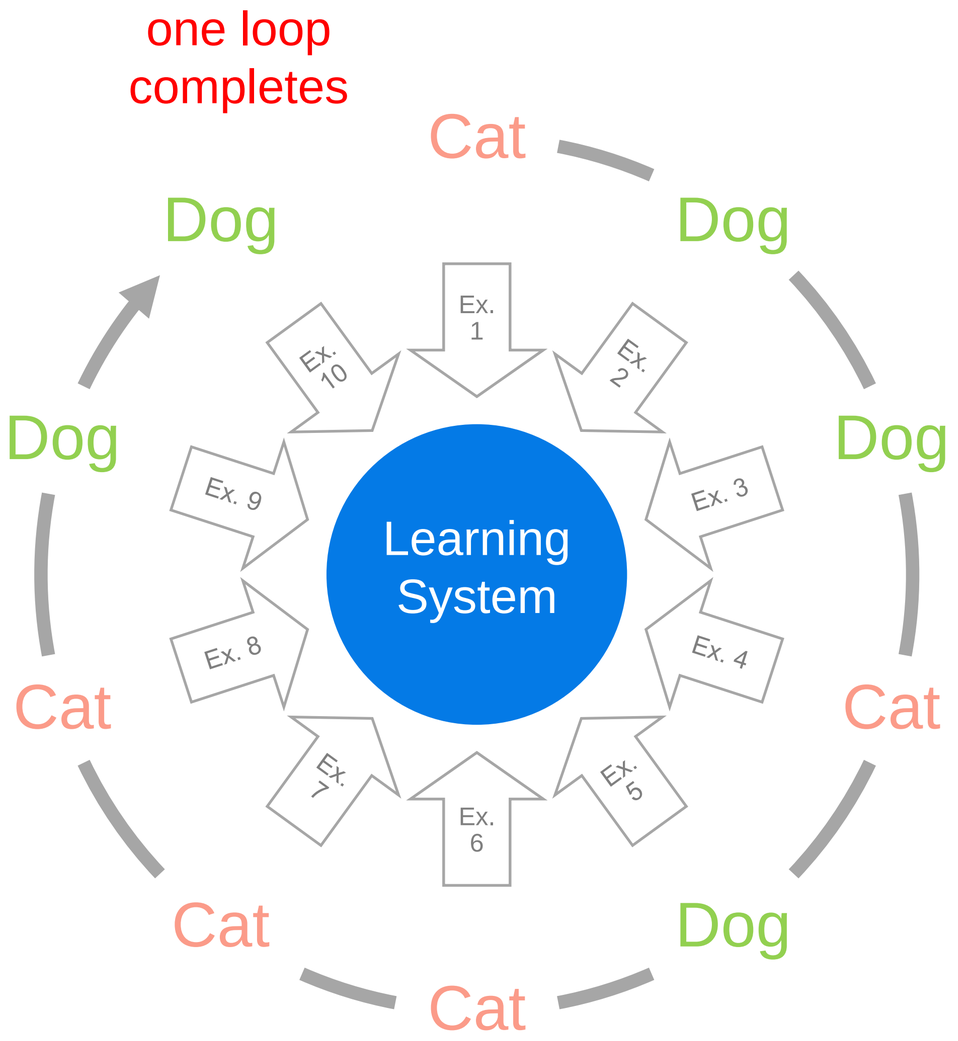
#	Inputs		Target
	Height (cm)	Weight (kg)	Animal
Ex. 1	23.2	3.2	Cat
Ex. 2	70.9	19.5	Dog
Ex. 3	60.5	18.51	Dog
Ex. 4	24.5	4.6	Cat
Ex. 5	110.0	35.83	Dog
Ex. 6	23.8	3.7	Cat
Ex. 7	25.8	4.5	Cat
Ex. 8	24.7	4.9	Cat
Ex. 9	85.8	25.6	Dog
Ex. 10	78.8	20.33	Dog



Data (labeled)

Discrete labeled data

#	Inputs		Target
	Height (cm)	Weight (kg)	Animal
Ex. 1	23.2	3.2	Cat
Ex. 2	70.9	19.5	Dog
Ex. 3	60.5	18.51	Dog
Ex. 4	24.5	4.6	Cat
Ex. 5	110.0	35.83	Dog
Ex. 6	23.8	3.7	Cat
Ex. 7	25.8	4.5	Cat
Ex. 8	24.7	4.9	Cat
Ex. 9	85.8	25.6	Dog
Ex. 10	78.8	20.33	Dog



Data (labeled)

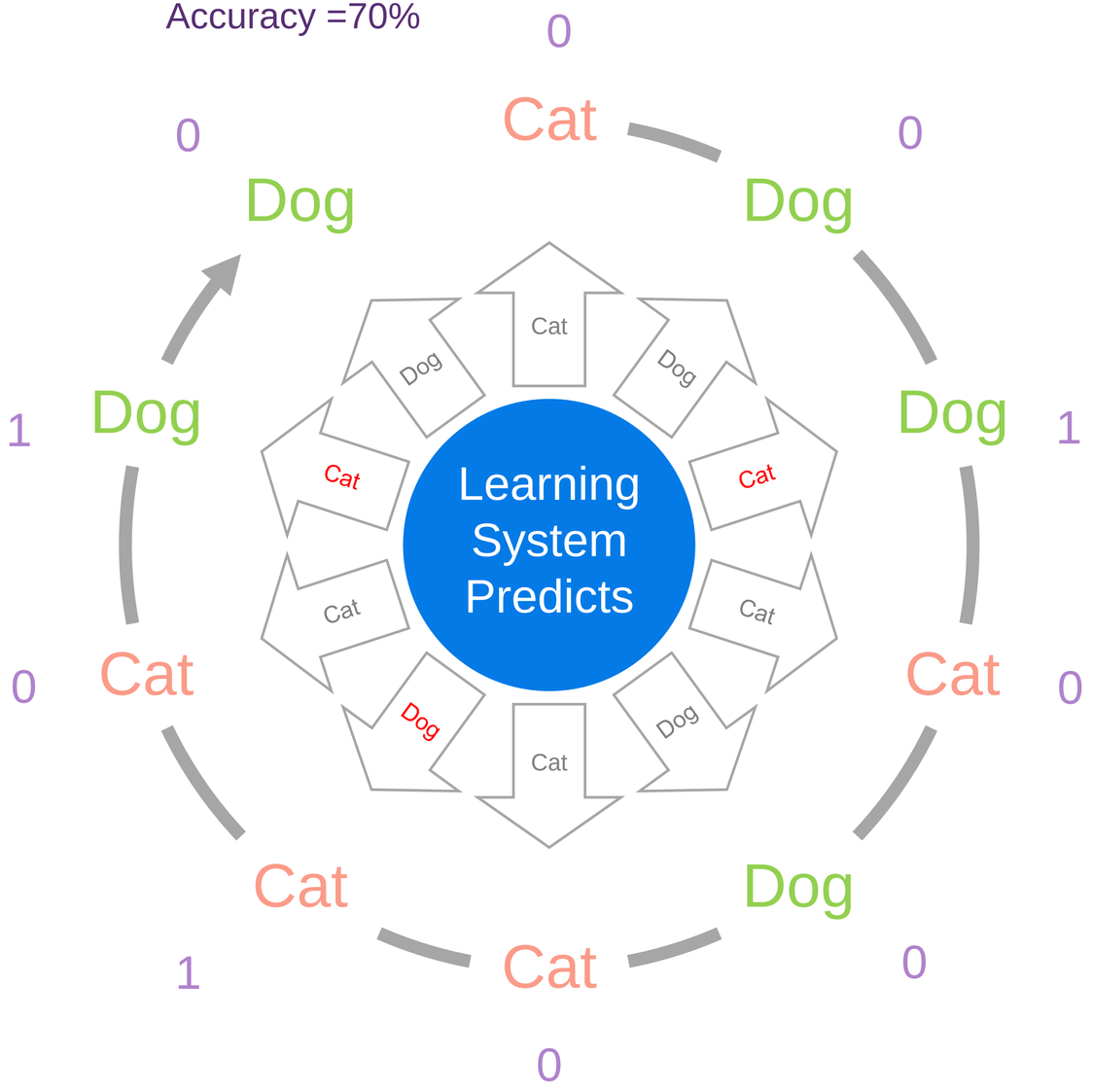
Discrete labeled data

#	Inputs		Target	Predicted	
	Height (cm)	Weight (kg)	Animal		Error
Ex. 1	23.2	3.2	Cat	Cat	0
Ex. 2	70.9	19.5	Dog	Dog	0
Ex. 3	60.5	18.51	Dog	Cat	0
Ex. 4	24.5	4.6	Cat	Cat	1
Ex. 5	110.0	35.83	Dog	Dog	0
Ex. 6	23.8	3.7	Cat	Cat	0
Ex. 7	25.8	4.5	Cat	Dog	1
Ex. 8	24.7	4.9	Cat	Cat	0
Ex. 9	85.8	25.6	Dog	Cat	1
Ex. 10	78.8	20.33	Dog	Dog	0

Loop 1

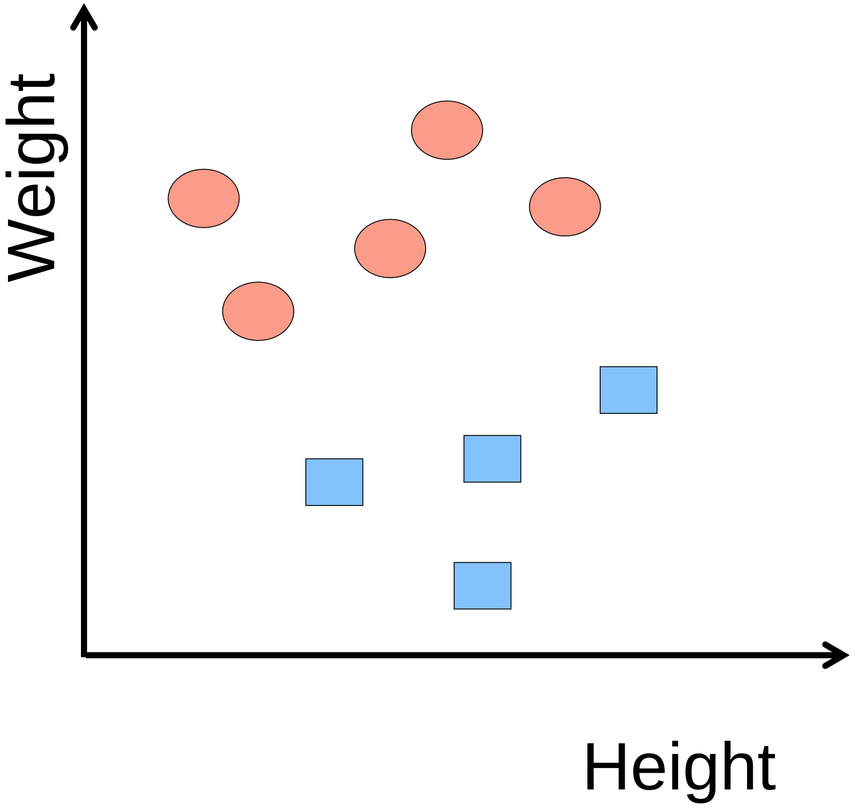
Error = 3/10,

Accuracy =70%

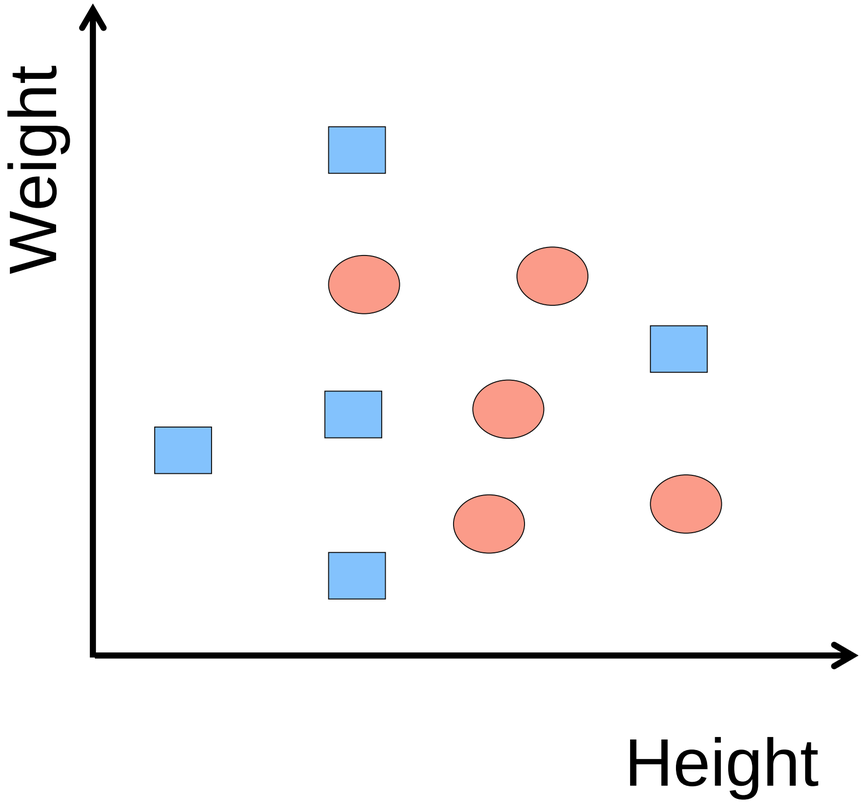


Classification

Linear classification

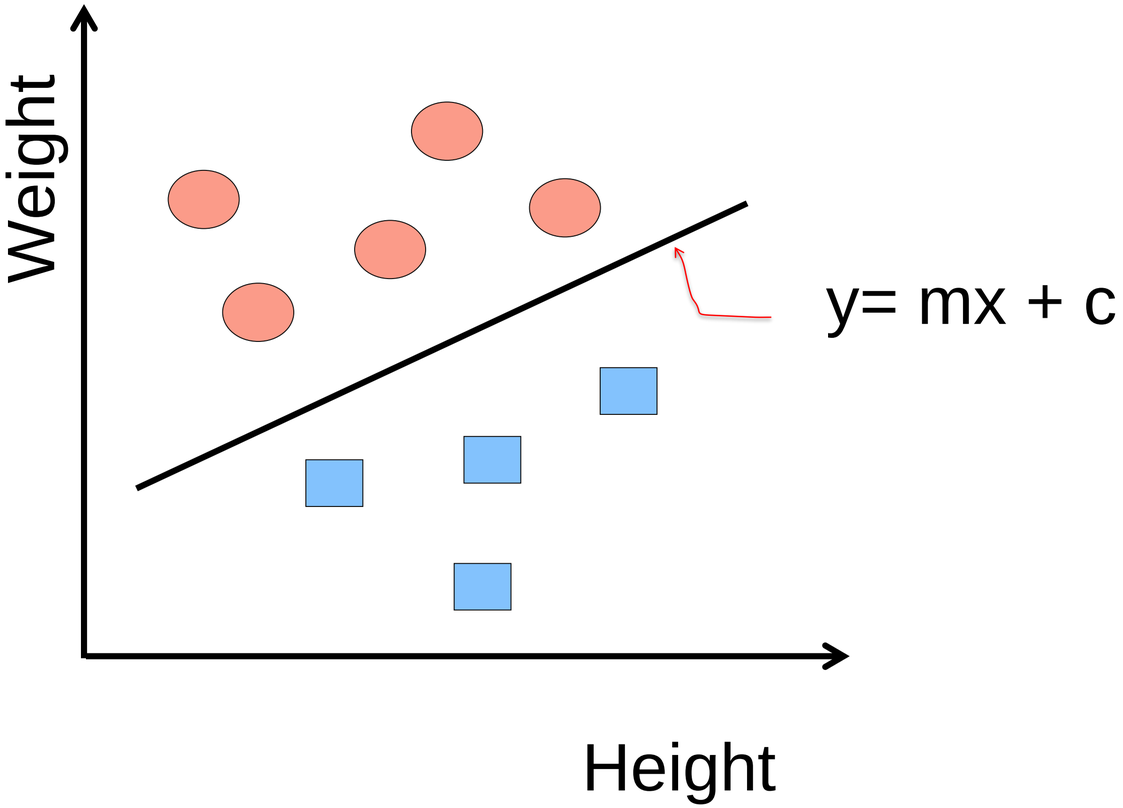


Non-linear classification

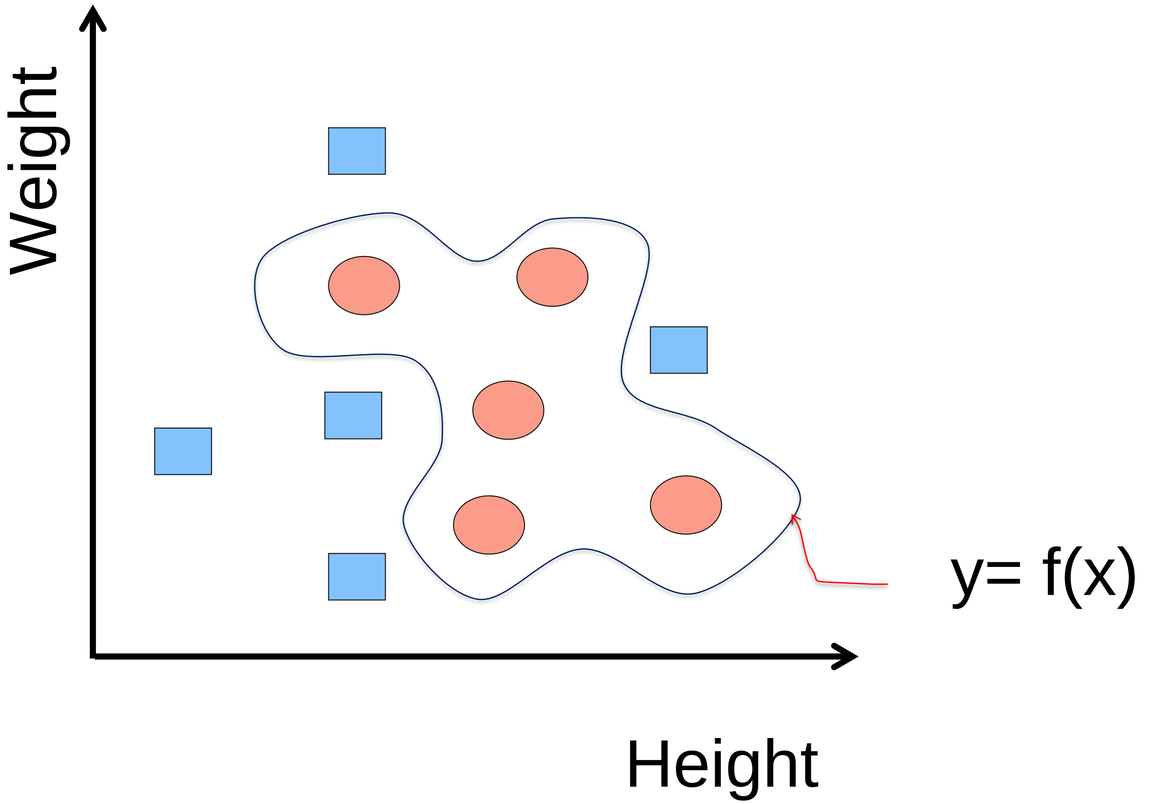


Classification

Linear classification



Non-linear classification



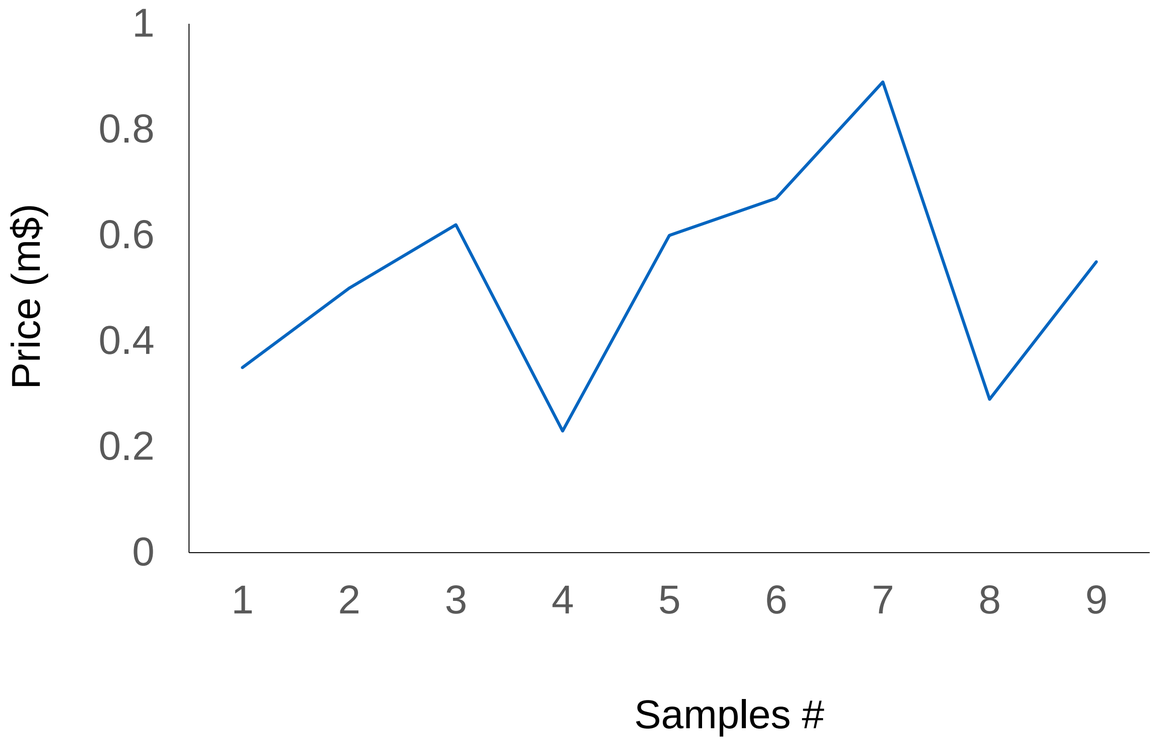
What is f(x)?

Data (labeled)

Target has continuous value

Continuous labeled data

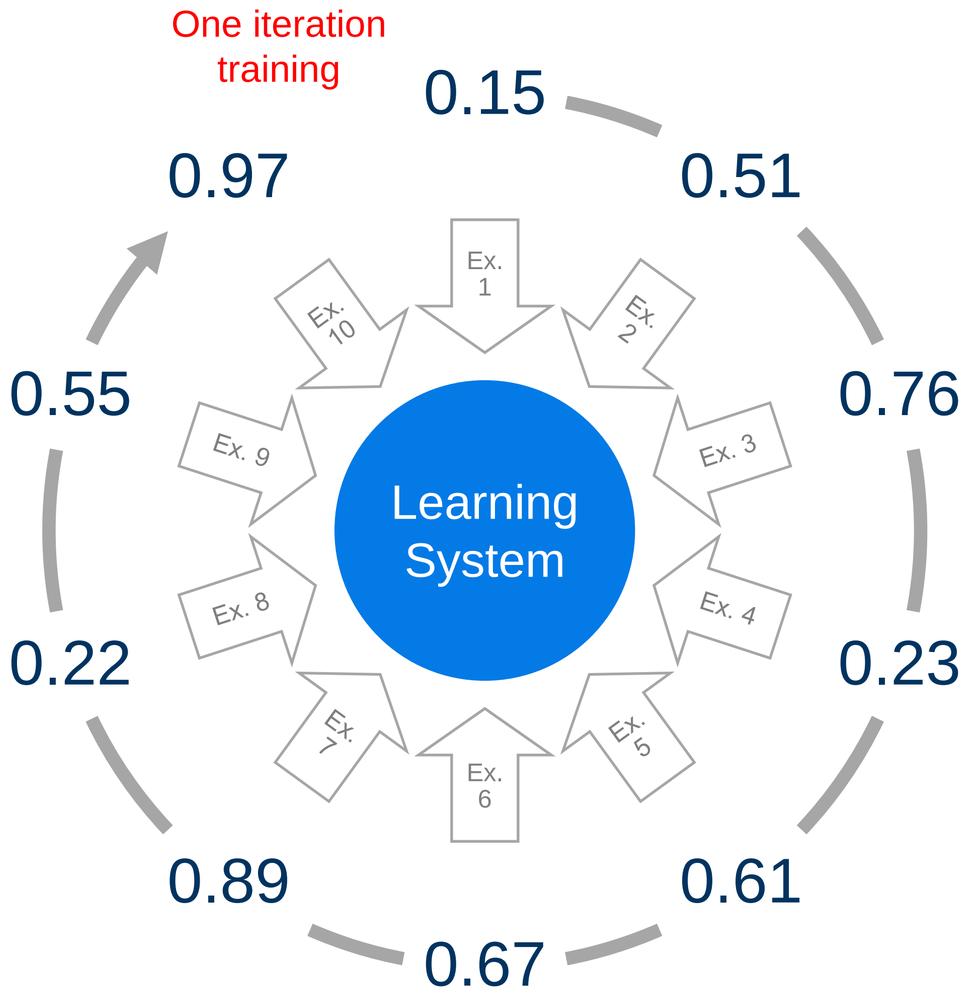
#	Inputs				Target
	Area	Distance	Location	Interior	Price (m$)
1	76.85	17.27	0.22	34.63	0.15
2	76.97	19.54	0.22	34.64	0.5
3	77.10	18.51	0.22	34.64	0.76
4	85.28	46.09	0.22	34.61	0.23
5	85.42	35.83	0.22	34.61	0.6
6	88.02	2.59	0.22	34.63	0.67
7	77.25	6.34	0.22	34.65	0.89
8	77.49	6.98	0.22	34.63	0.2
9	85.81	12.18	0.22	34.61	0.55



Data (labeled)

Discrete labeled data

#	Inputs				Target
	Area	Distance	Location	Interior	Price (m$)
1	76.85	17.27	0.22	34.63	1.5
2	76.97	19.54	0.22	34.64	5
3	77.10	18.51	0.22	34.64	7.6
4	85.28	46.09	0.22	34.61	2.3
5	85.42	35.83	0.22	34.61	6.1
6	88.02	2.59	0.22	34.63	6.7
7	77.25	6.34	0.22	34.65	8.9
8	77.49	6.98	0.22	34.63	2.2
9	85.81	12.18	0.22	34.61	5.5
10	88.02	2.59	0.22	34.63	9.7



Data (labeled)

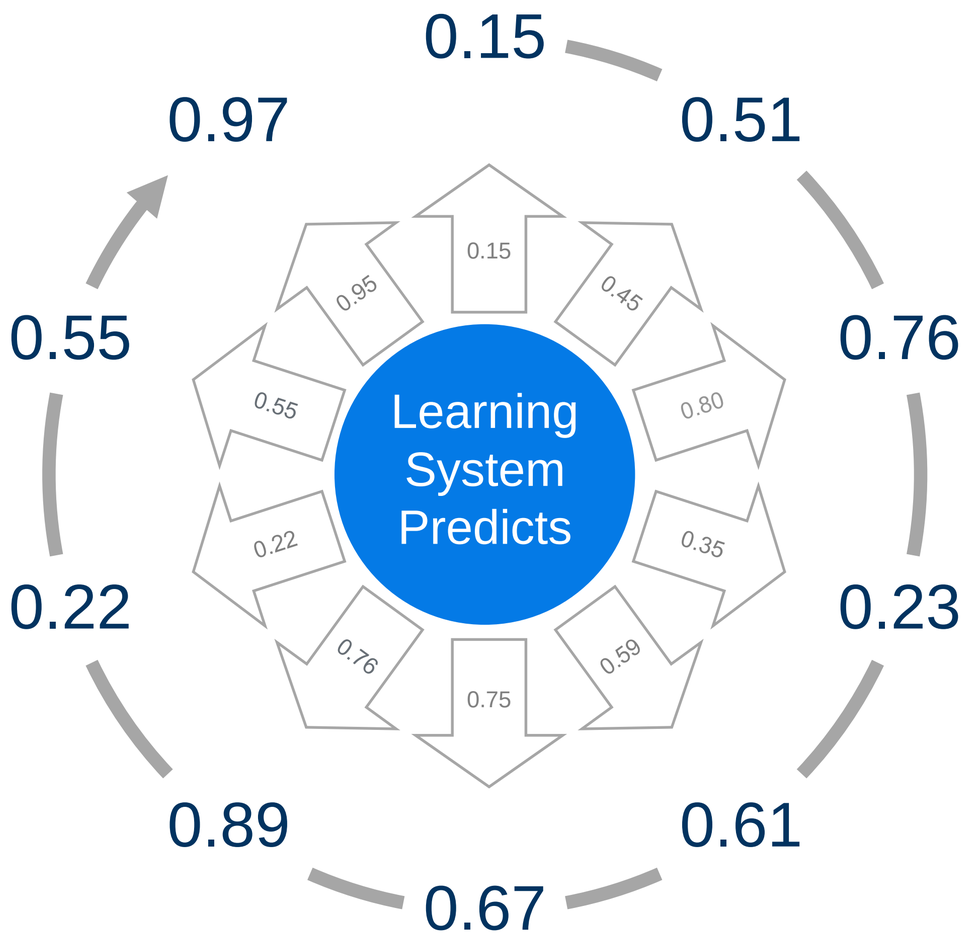
Loop 1

MSE Error = 0.43,

Correlation = 0.97

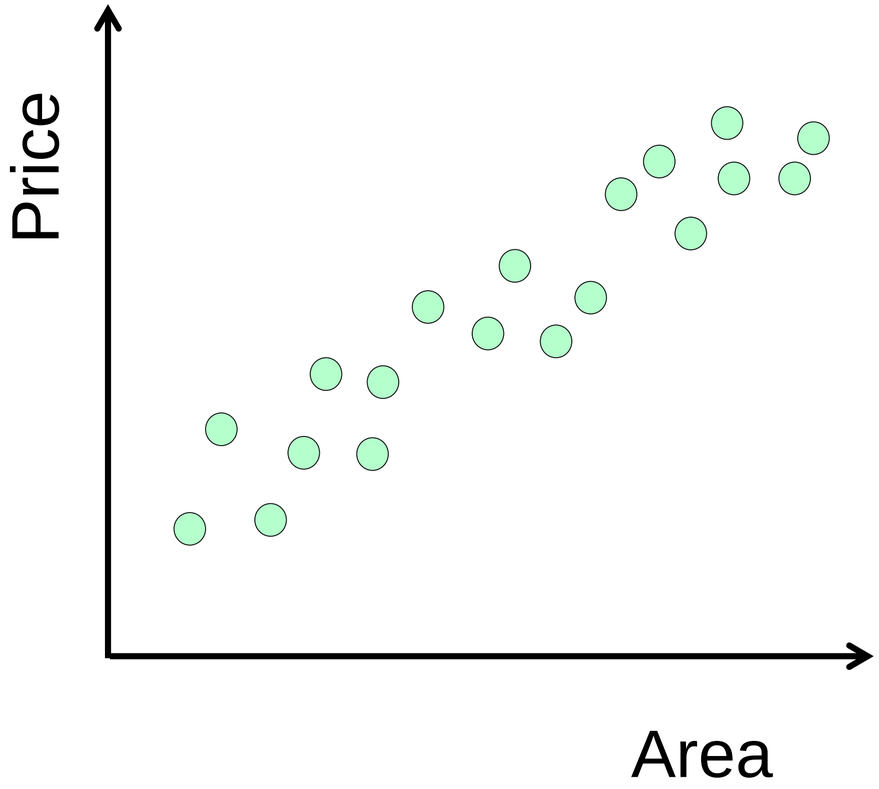
Discrete labeled data

#	Inputs				Target	Predicted	Error
	Area	Distance	Location	Interior	Price (m$)		
1	76.85	17.27	0.22	34.63	1.5	1.5	0
2	76.97	19.54	0.22	34.64	5	4.5	0.5
3	77.10	18.51	0.22	34.64	7.6	8	-0.4
4	85.28	46.09	0.22	34.61	2.3	3.5	-1.2
5	85.42	35.83	0.22	34.61	6.1	5.9	0.2
6	88.02	2.59	0.22	34.63	6.7	7.5	-0.8
7	77.25	6.34	0.22	34.65	8.9	7.6	1.3
8	77.49	6.98	0.22	34.63	2.2	2.2	0
9	85.81	12.18	0.22	34.61	5.5	5.5	0
10	88.02	2.59	0.22	34.63	9.7	9.5	0.2

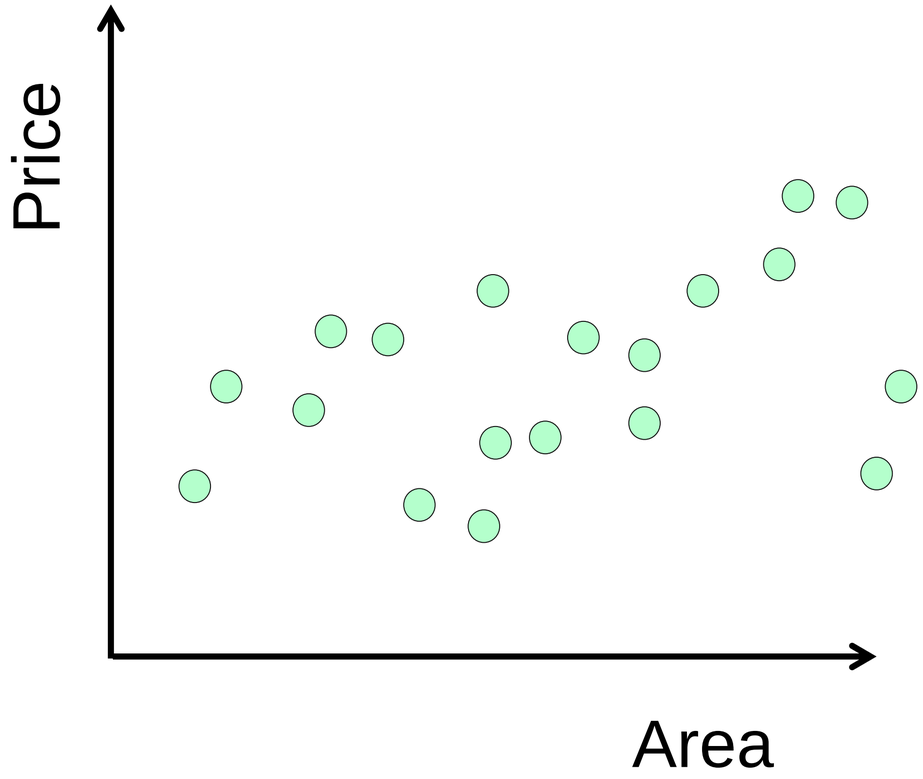


Regression

Linear regression

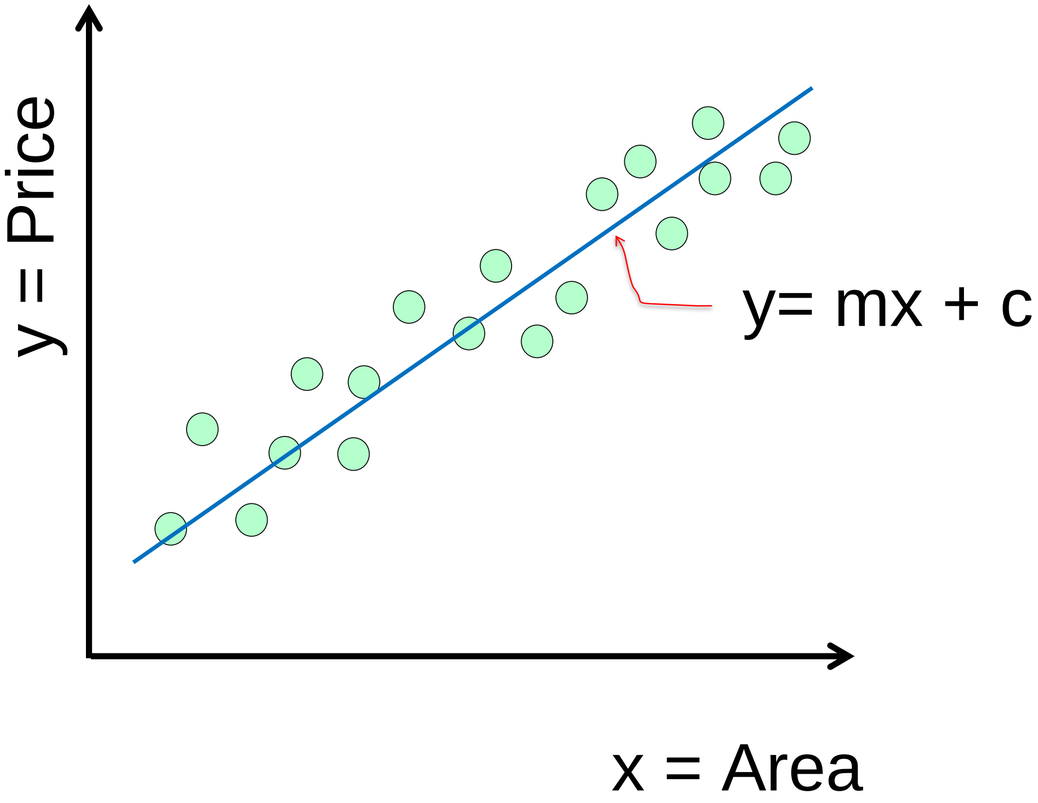


Non-linear regression

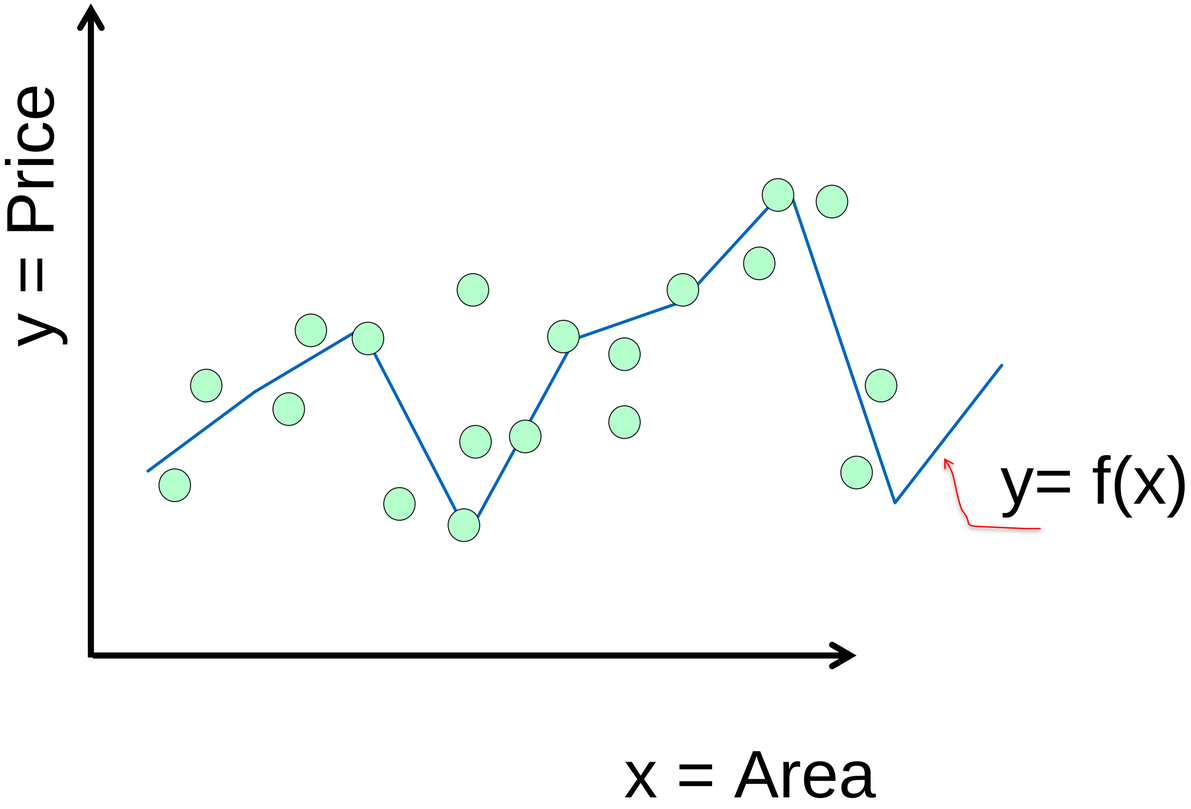


Regression

Linear regression



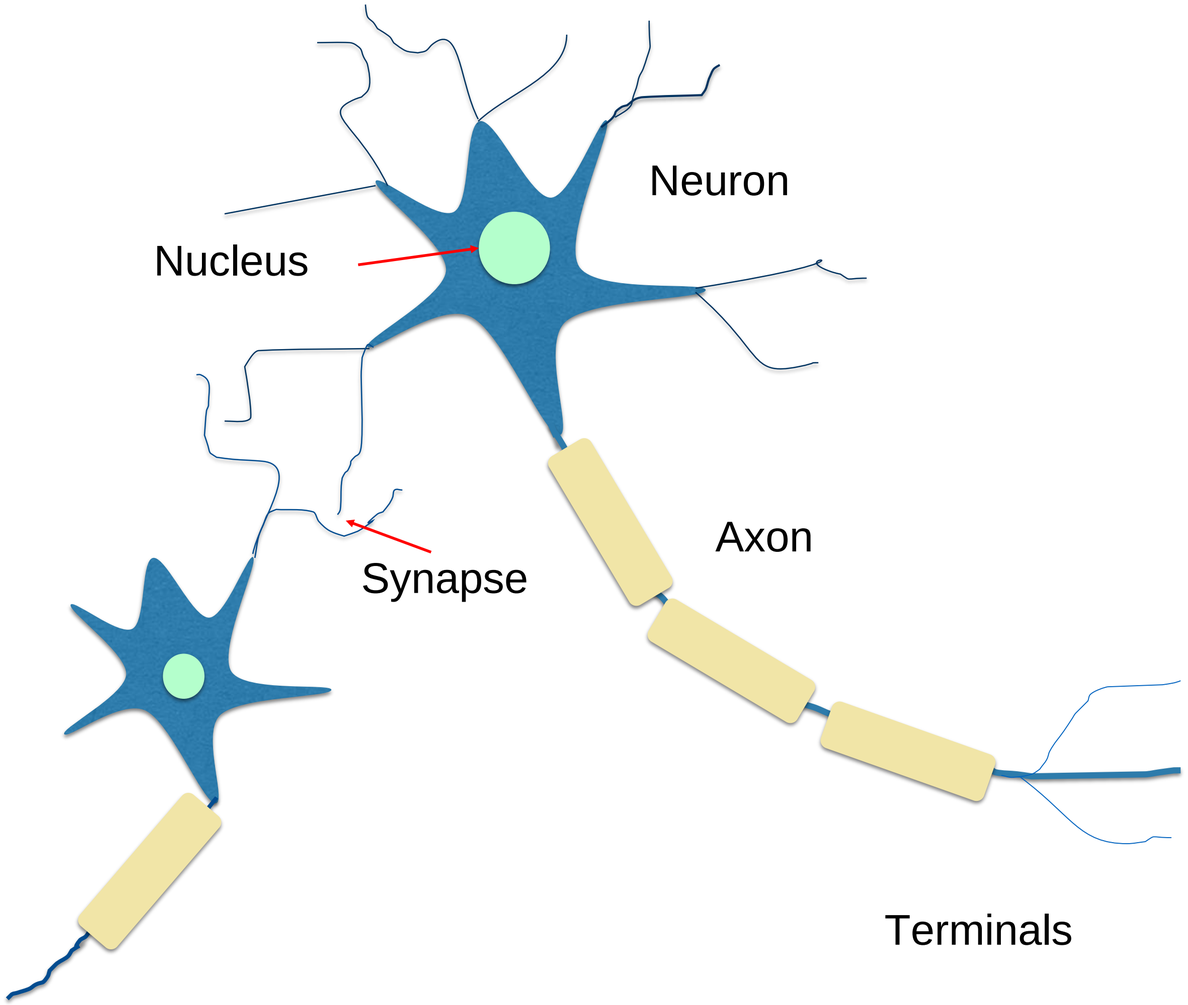
Non-linear regression



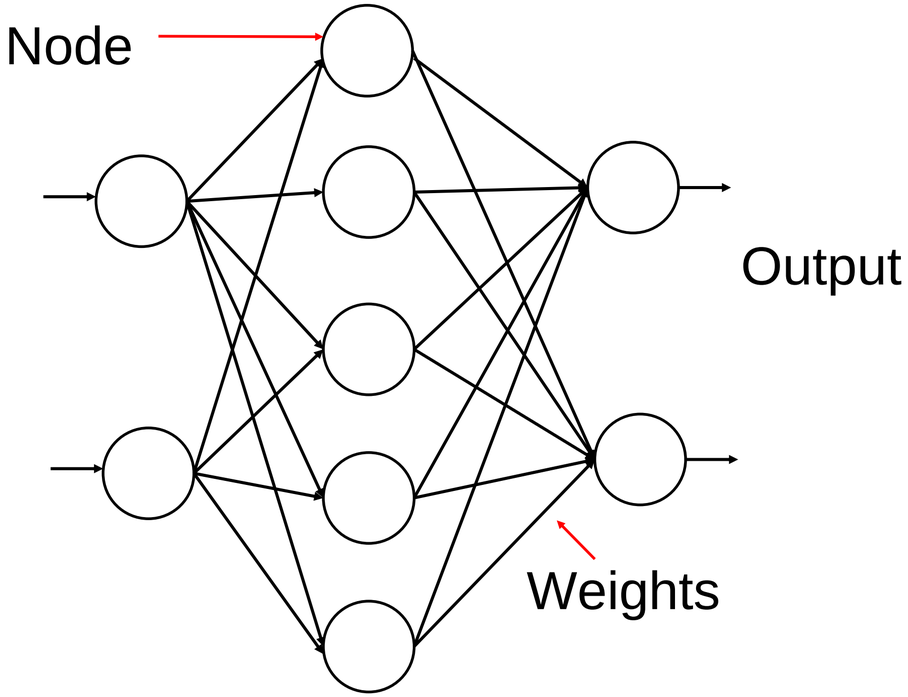
What is f(x)?

Neural Network

Biological neural network



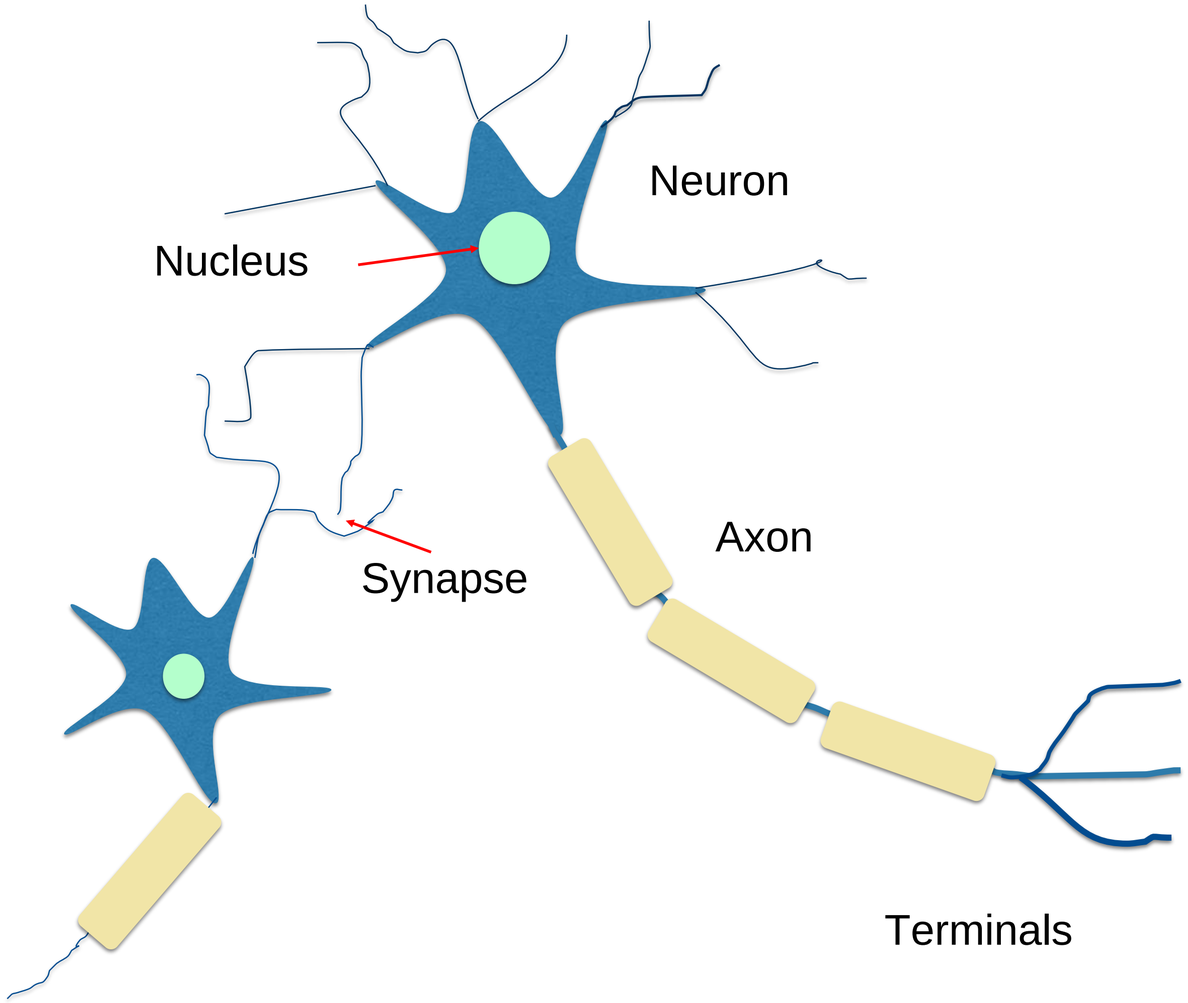
Artificial neural network



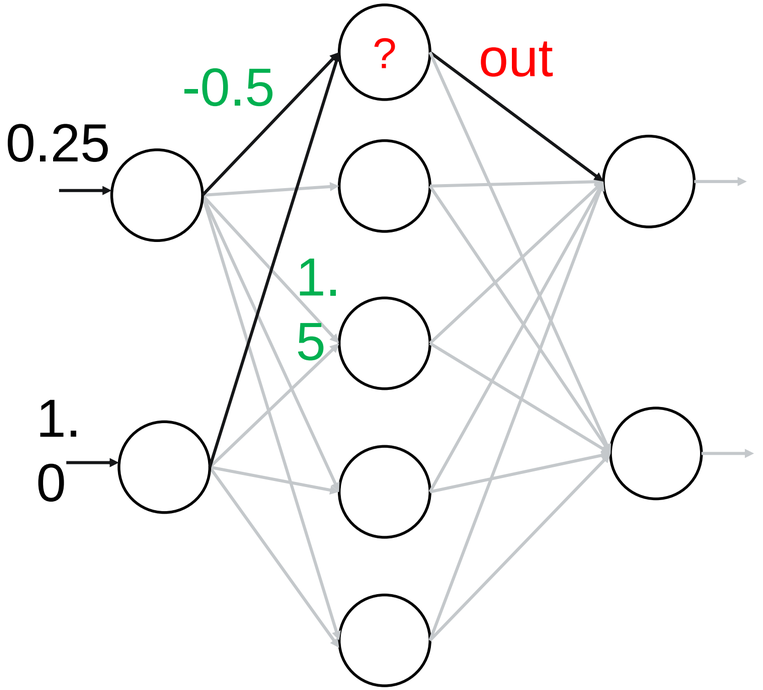
Neuron -> Node
Nucleus -> Squashing function
Synapse -> Weights
Terminals -> Output

Neural Network

Biological neural network



Artificial neural network

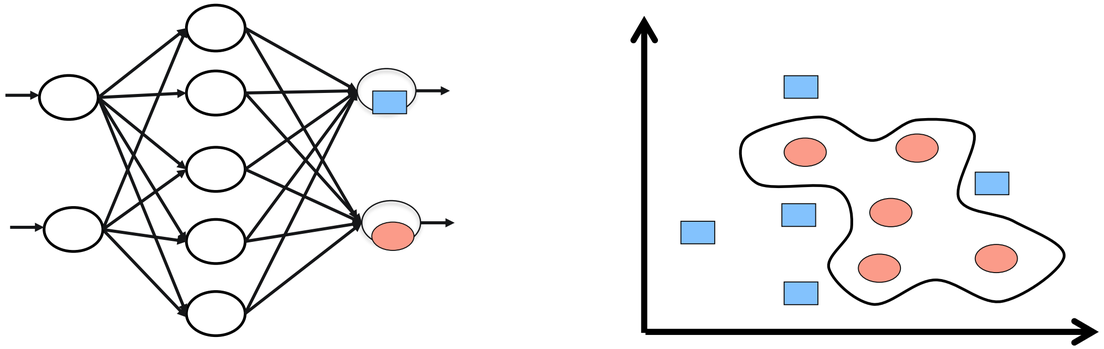


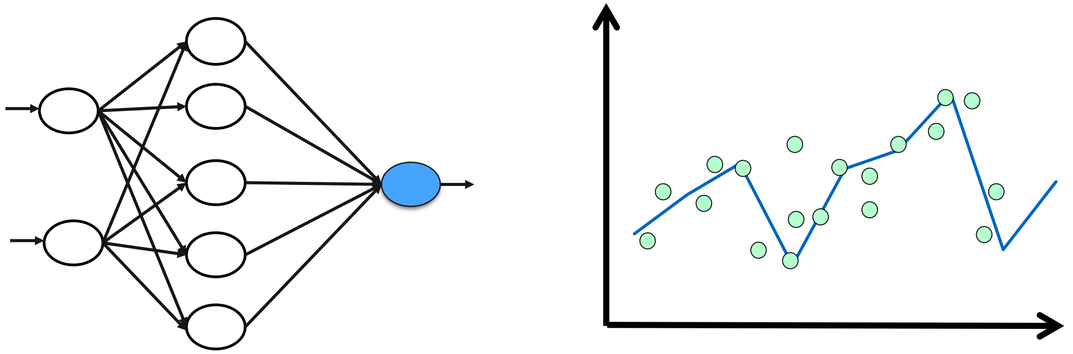
sum = (0.25x-0.5) + (1.0x1.5) = 1.625

$$\text{out} = \frac{1}{1+e^{1.525}} = 0.165$$ Squashed value

What is f(x)?

Neural Network → f(x)





Neural Network will give a function f(x) for which f(x) will separate two classes with a maximum accuracy

Neural Network will give a function f(x) for which f(x) will fit within the data points with less error.

Thank you!

Varun Ojha ojha@arch.ethz.ch

Danielle Griego griego@arch.ethz.ch