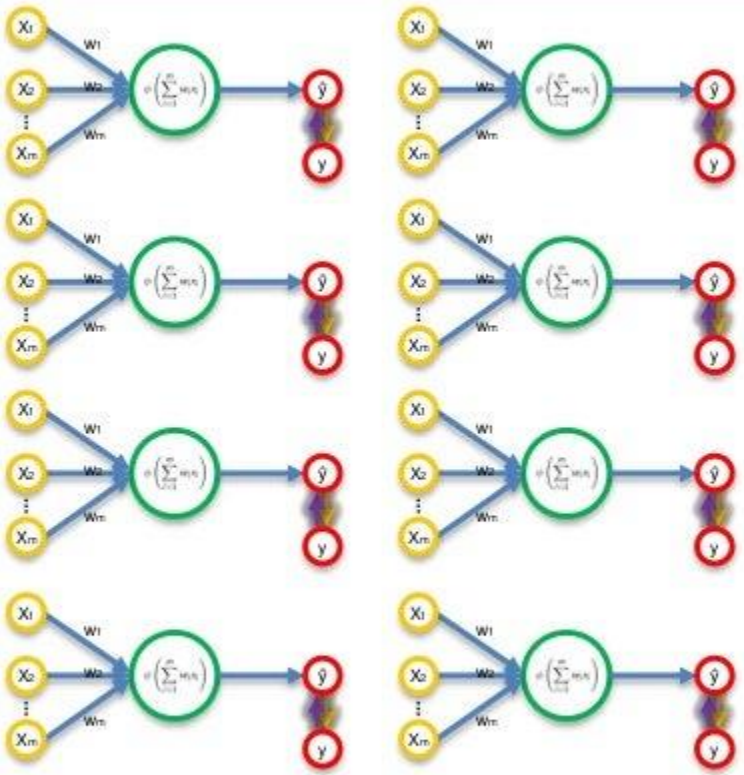
Gradient Descent

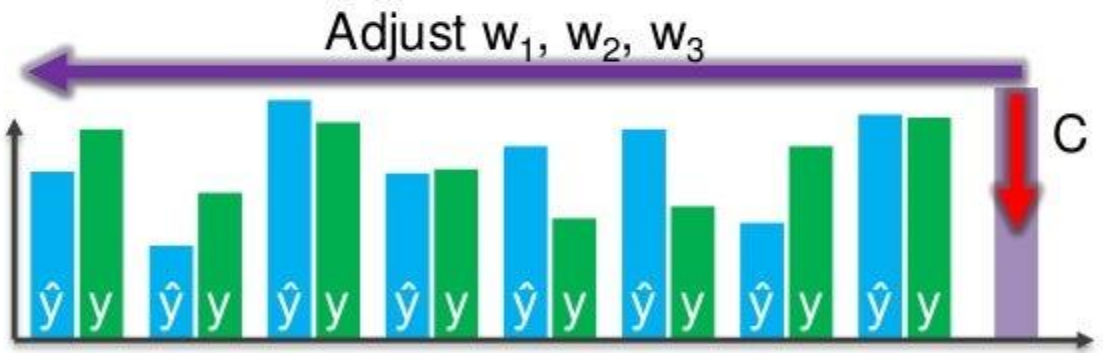
© SuperDataScience

Gradient Descent



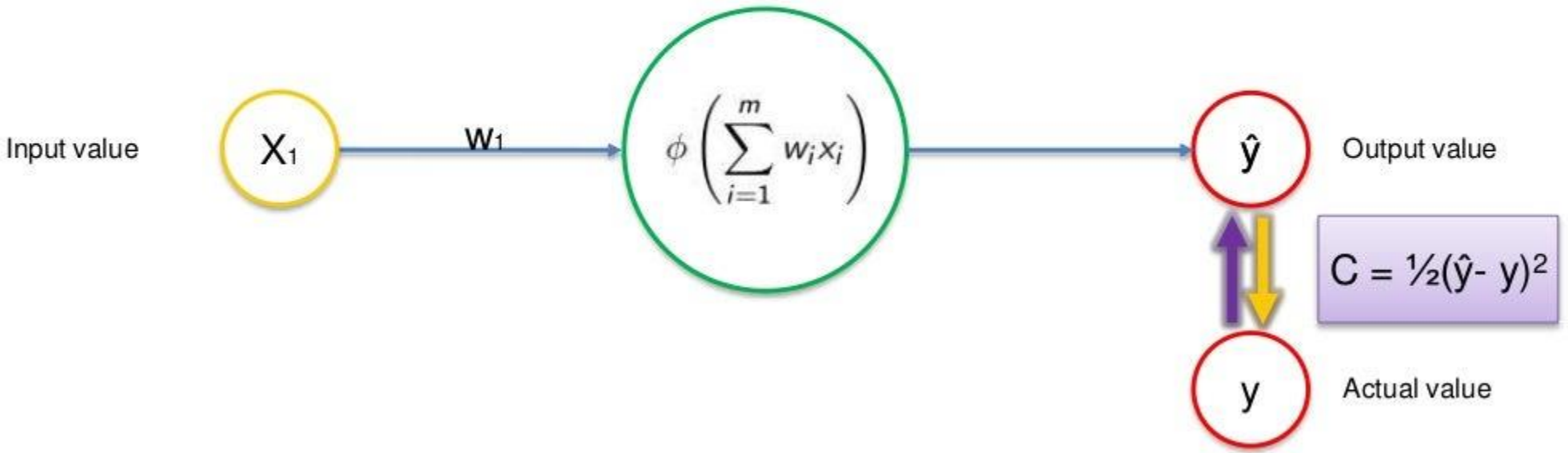
Row ID	Study Hrs	Sleep Hrs	Quiz	Exam
1	12	6	78%	93%
2	22	6.5	24%	68%
3	115	4	100%	95%
4	31	9	67%	75%
5	0	10	58%	51%
6	5	8	78%	60%
7	92	6	82%	89%
8	57	8	91%	97%

$$C = \sum \tfrac{1}{2}(\hat{y} - y)^2$$



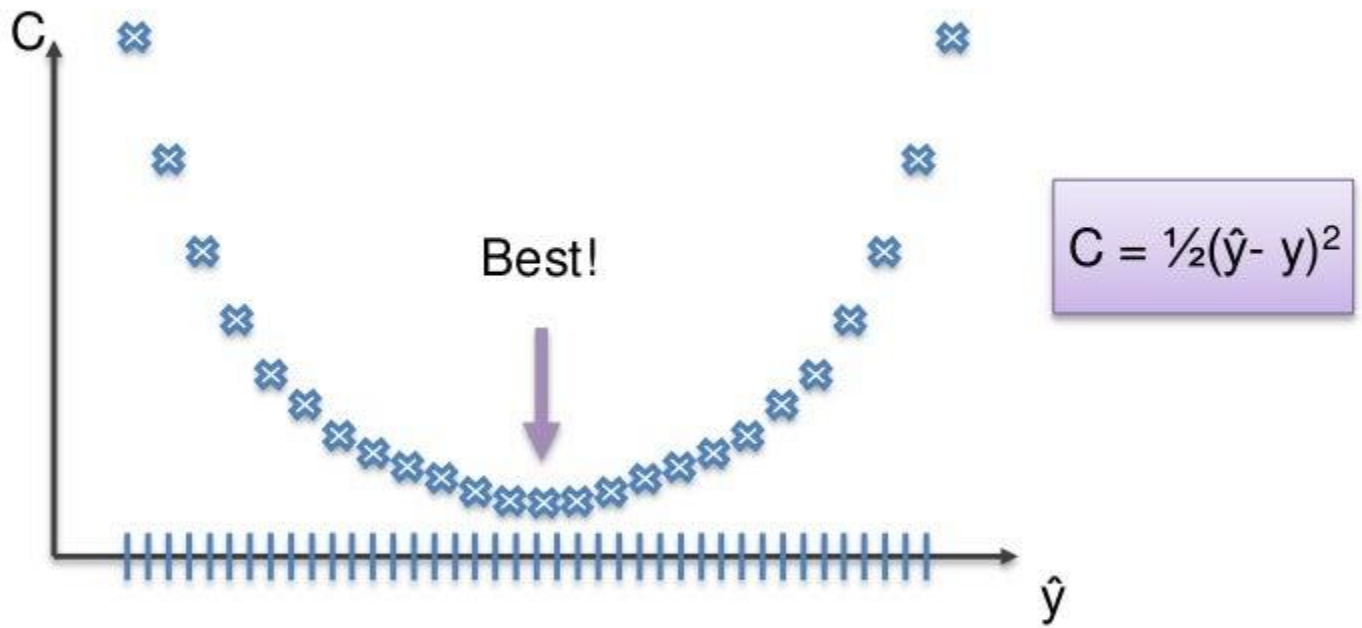
© SuperDataScience

Gradient Descent



© SuperDataScience

Gradient Descent

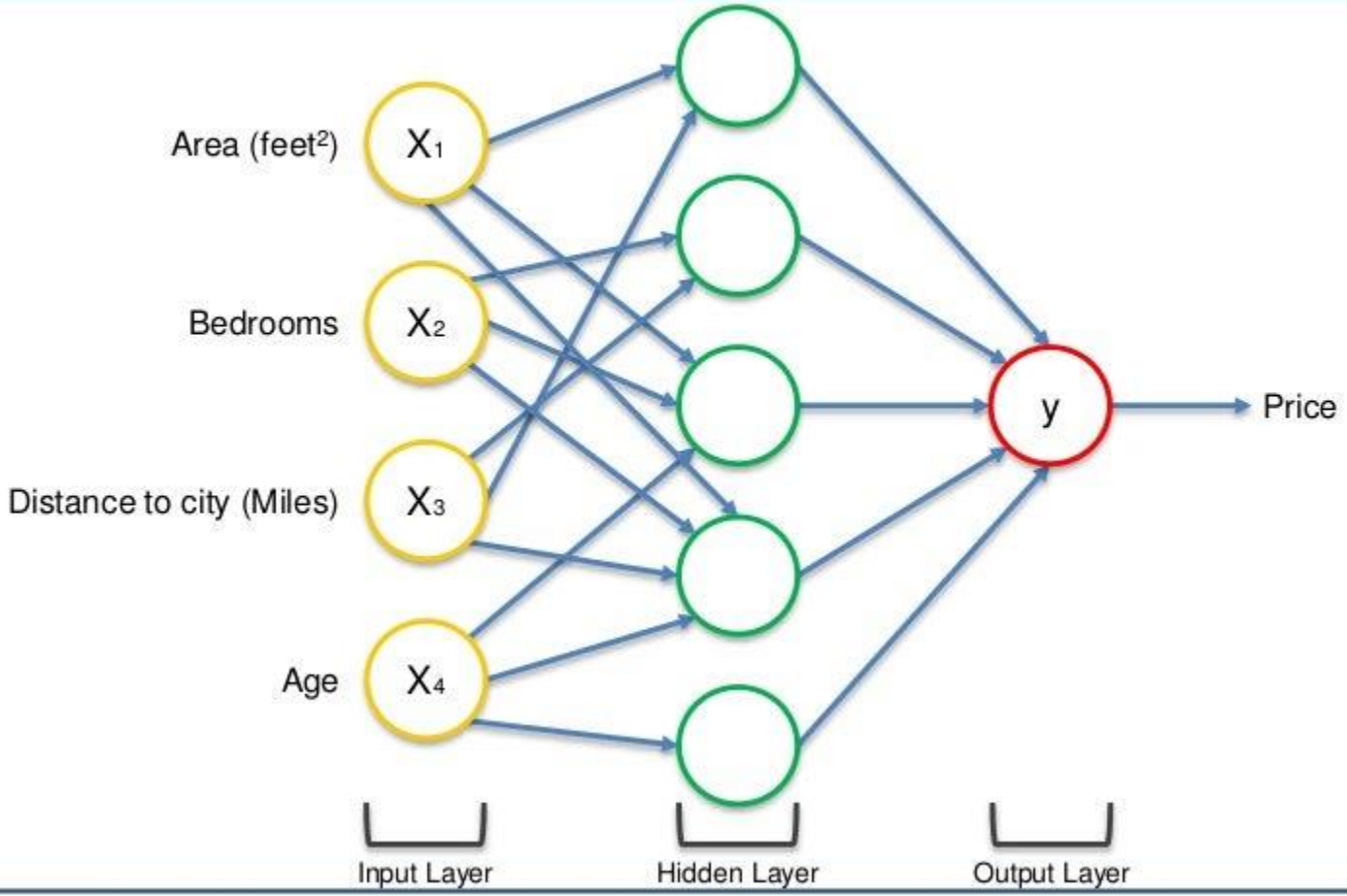


© SuperDataScience

Curse of Dimensionality

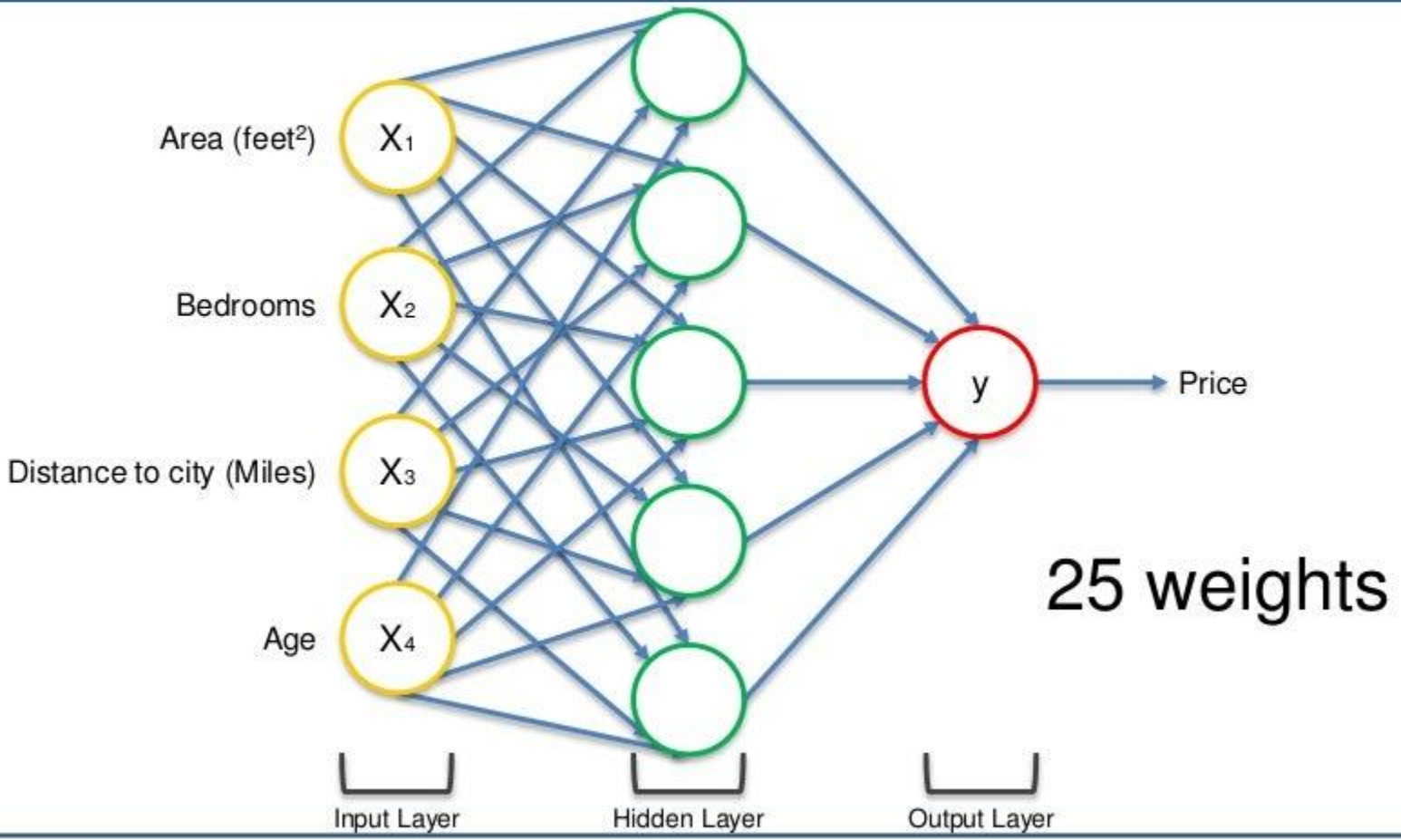
© SuperDataScience

Gradient Descent



© SuperDataScience

Gradient Descent



© SuperDataScience

Gradient Descent

$1{,}000 \times 1{,}000 \times \ldots \times 1{,}000 = 1{,}000^{25} = 10^{75}$ combinations

Sunway TaihuLight: World's fastest Super Computer

93 PFLOPS

93×10^{15}

$10^{75} / (93 \times 10^{15})$

$= 1.08 \times 10^{58}$ seconds

$= 3.42 \times 10^{50}$ years

© SuperDataScience

Gradient Descent

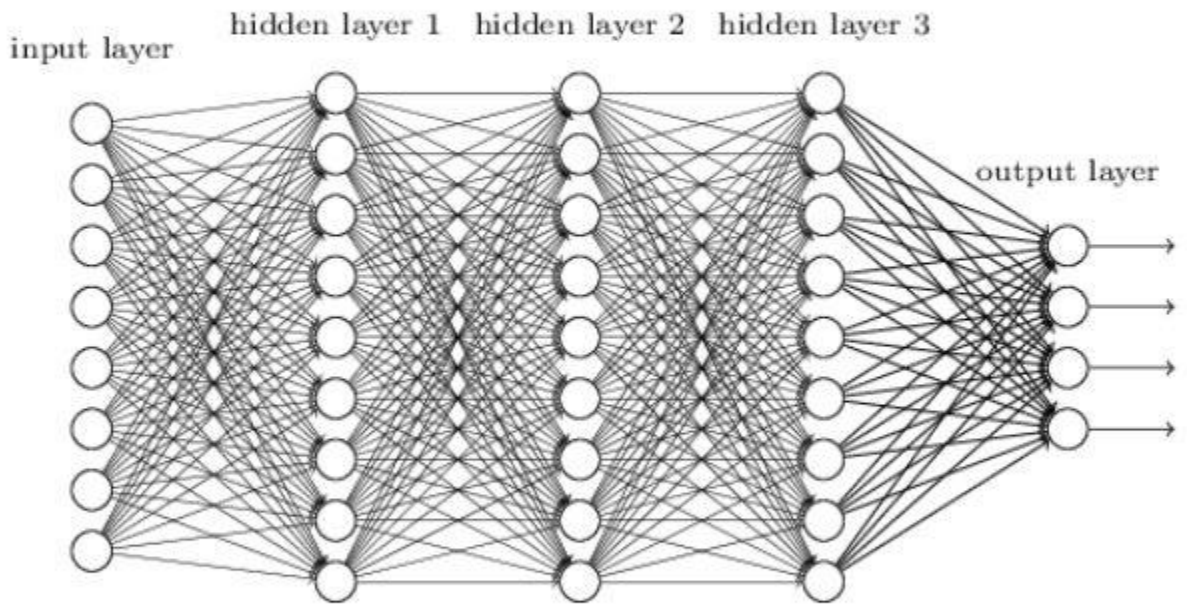


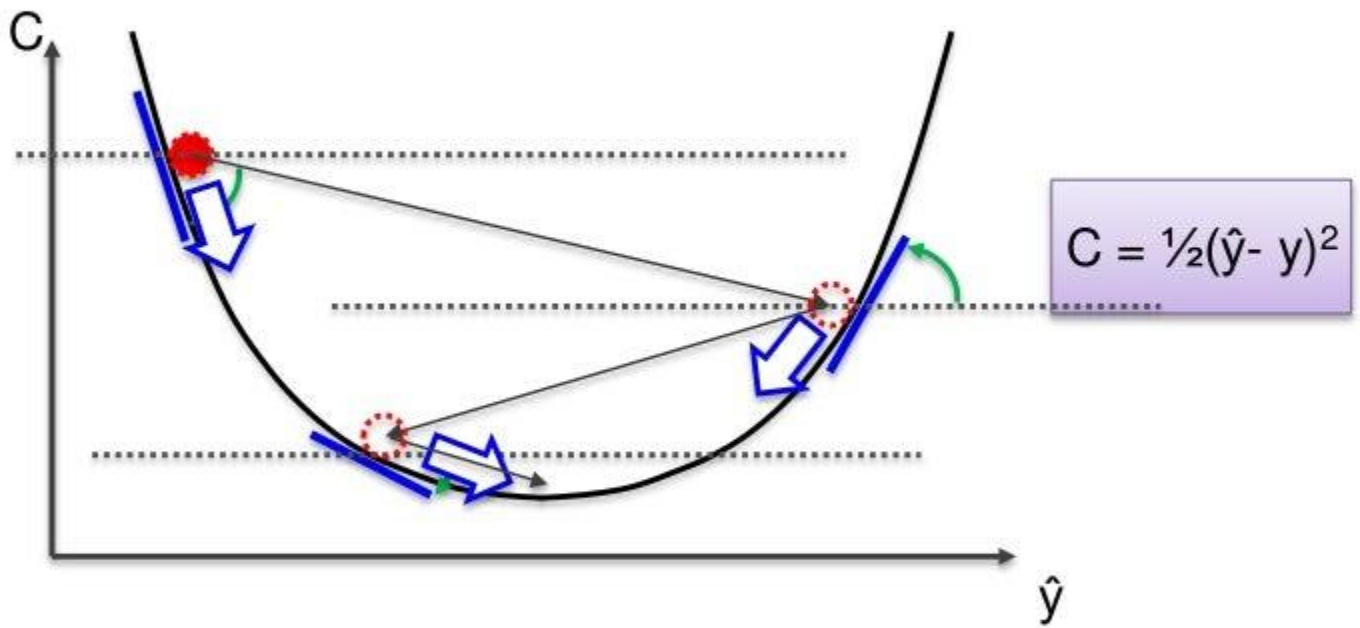
Image Source: neuralnetworksanddeeplearning.com

© SuperDataScience

Gradient Descent

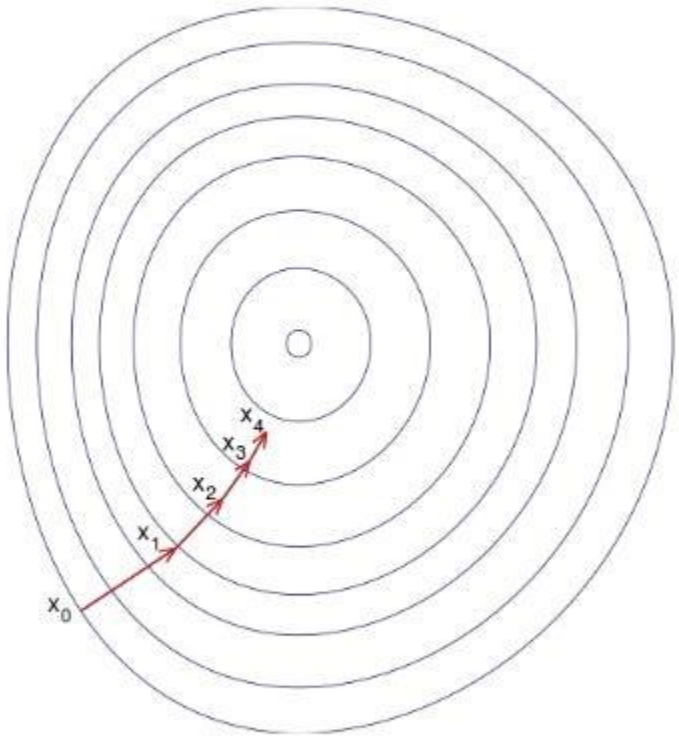
© SuperDataScience

Gradient Descent



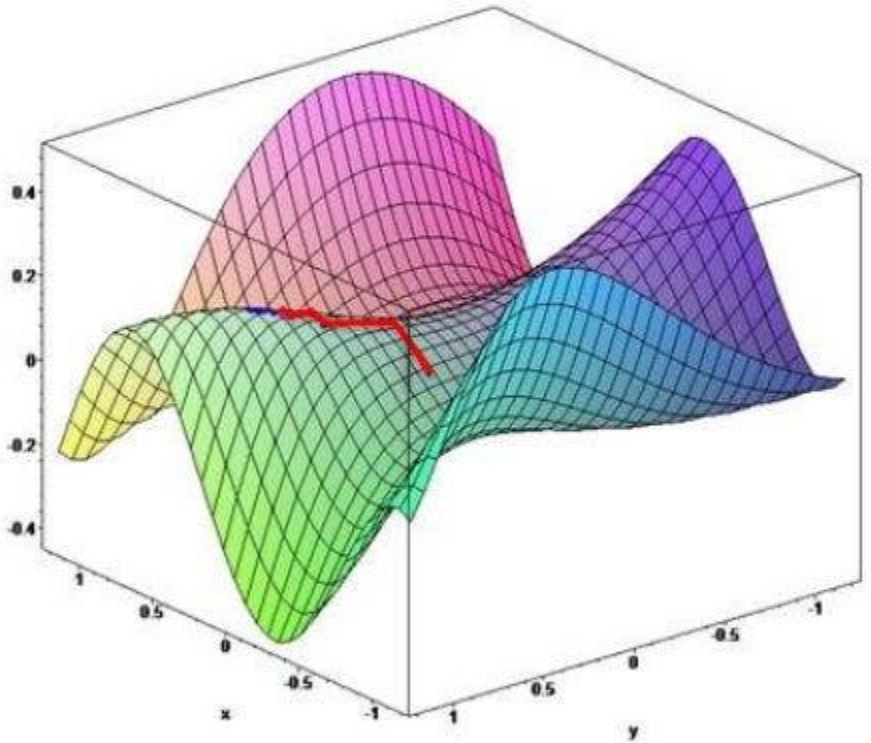
© SuperDataScience

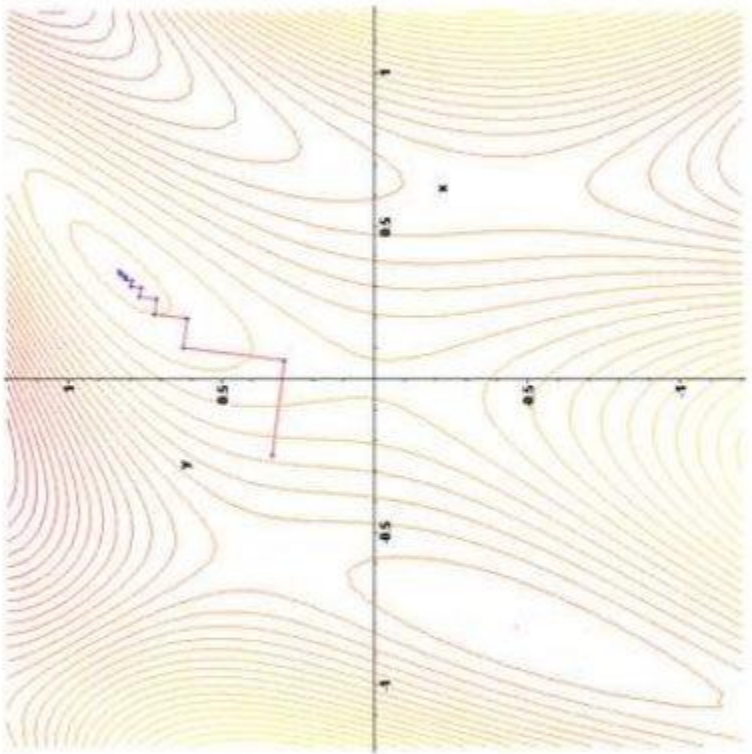
Gradient Descent



© SuperDataScience

Gradient Descent





© SuperDataScience