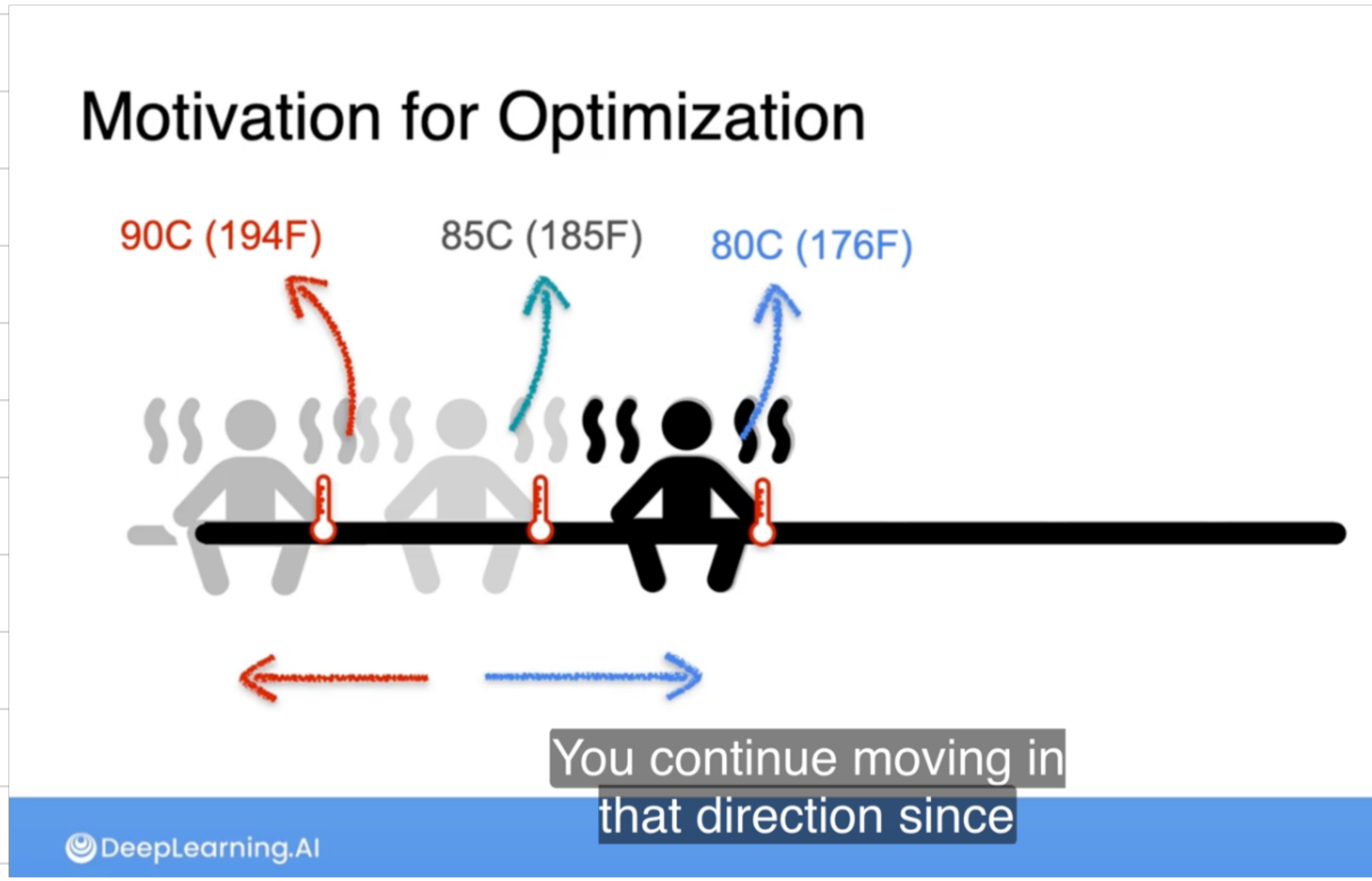
Day-82, Feb-20, 2025 (Falgun 8, 2081)

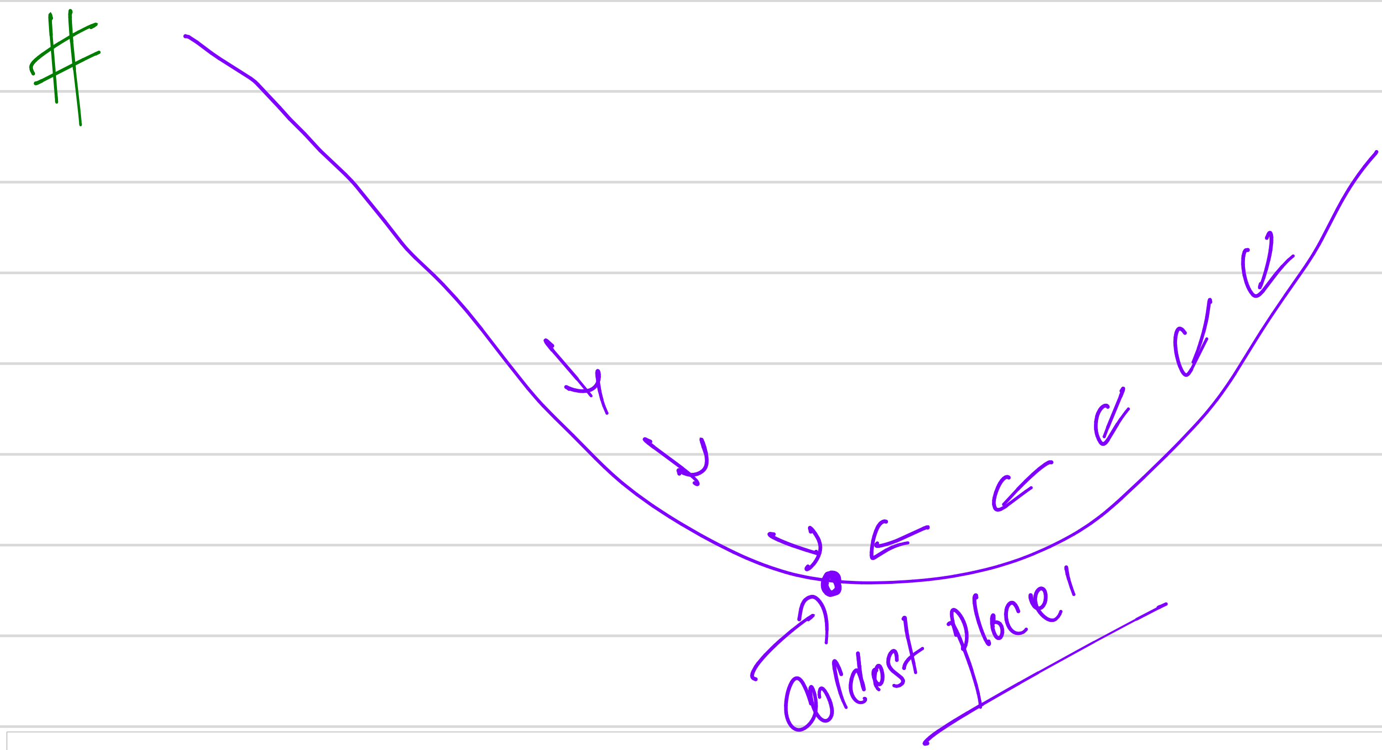
1. Optimization
2. Optimization of Squared loss → The one powerline problem
3. Optimization of Squared loss → The two Power line problem
4. Optimization of Squared loss → The three Powerline "
5. Optimization of Log-loss
6. Introduction to Tangent Planes
7. Partial Derivatives Part I and Part-II.

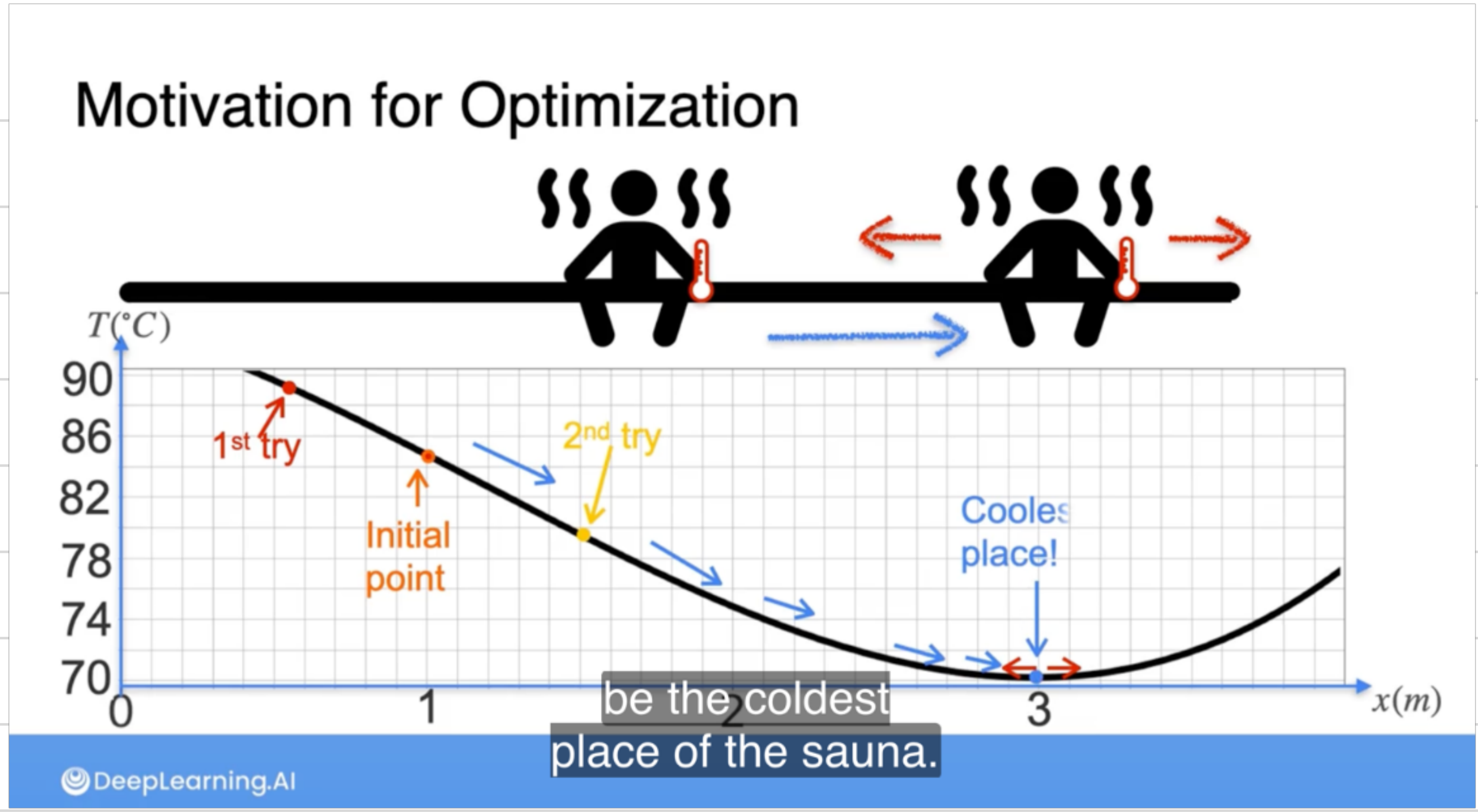
Introduction to optimization

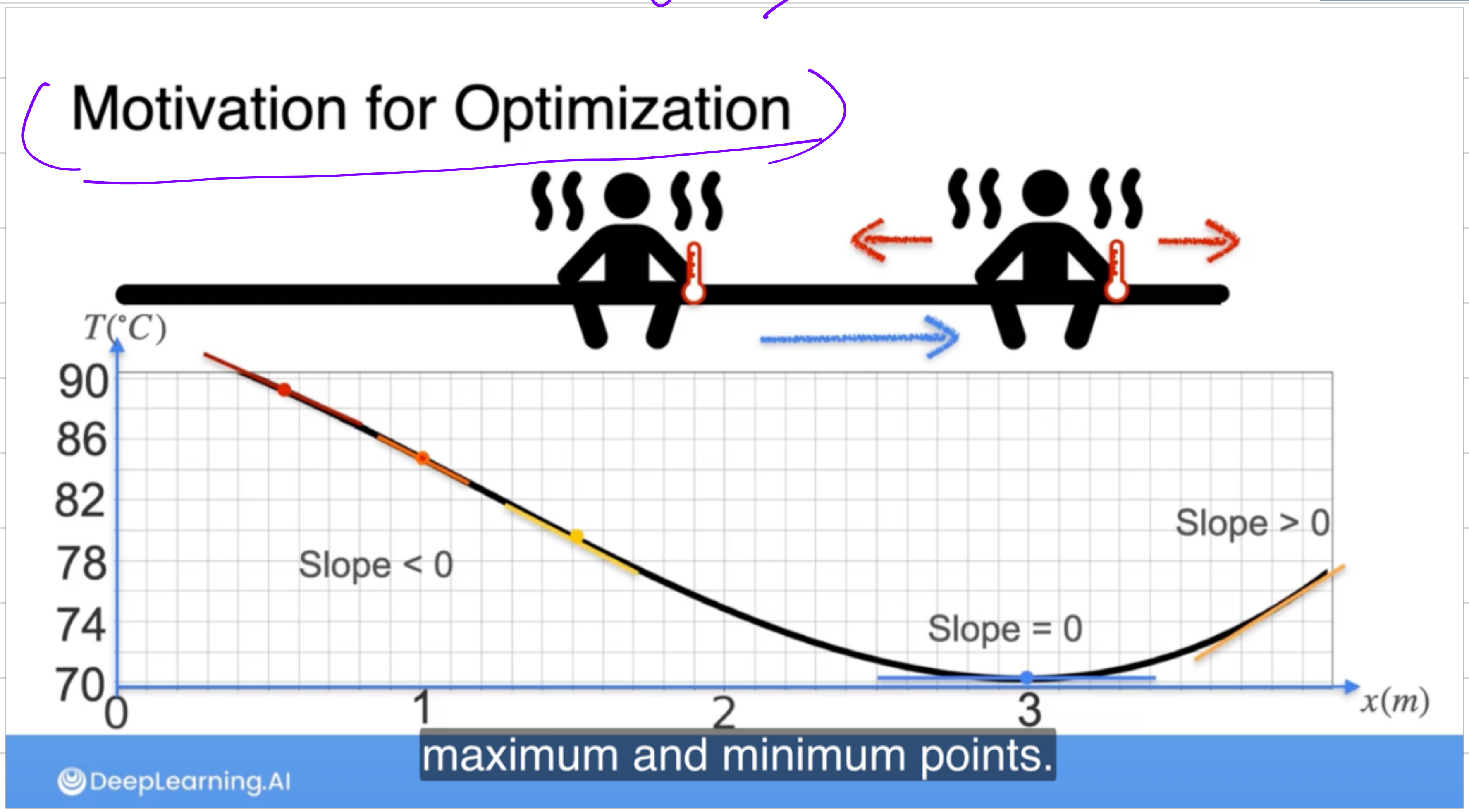
Optimization is when you want to find the maximum or the minimum value of a function. To find the best fit model we must be able to minimize an error function that tells us how far are we from the ideal model.

Optimization Problem trying to give or find the



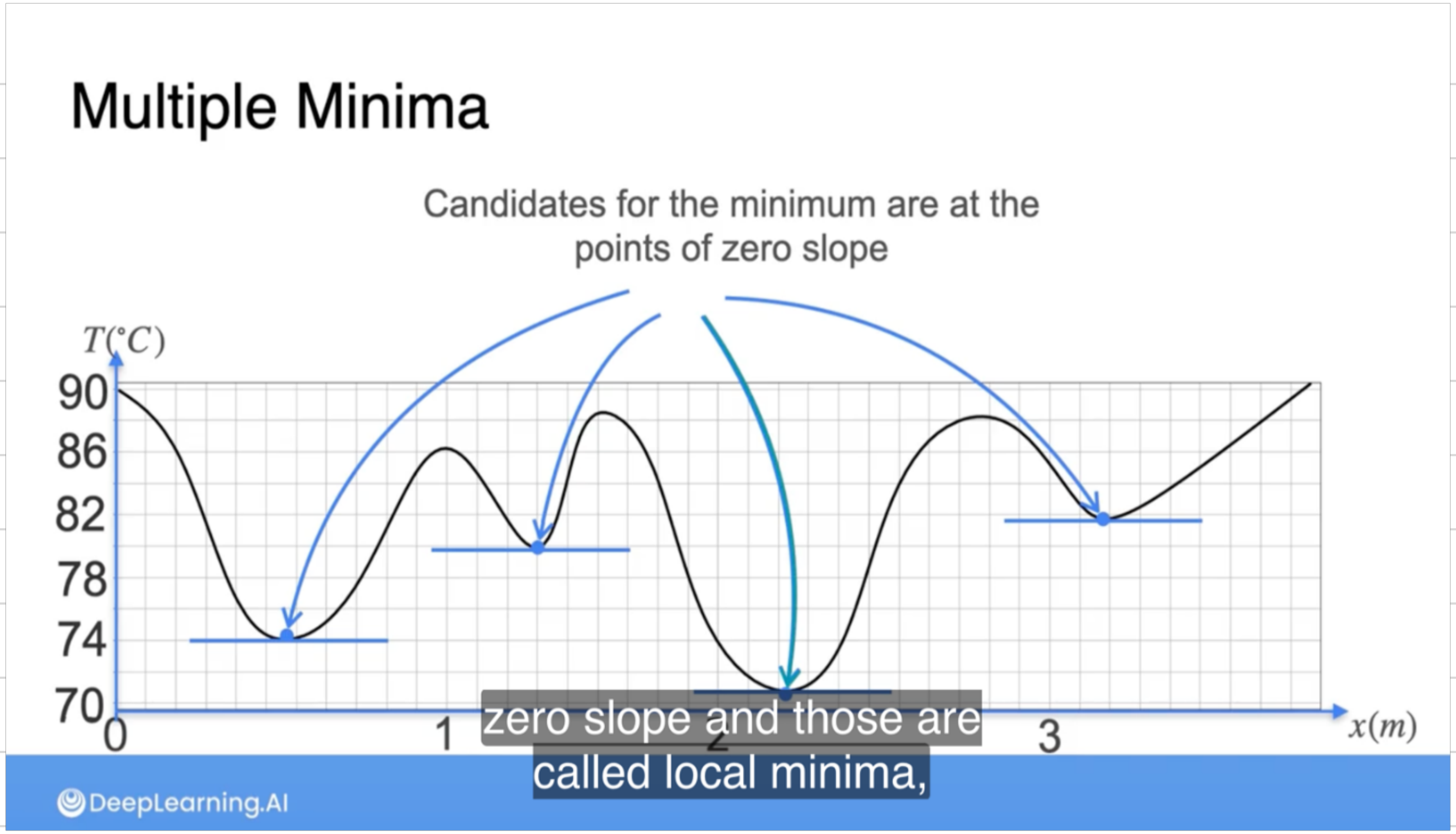






→ Concepts Similar to Hill Climbing.

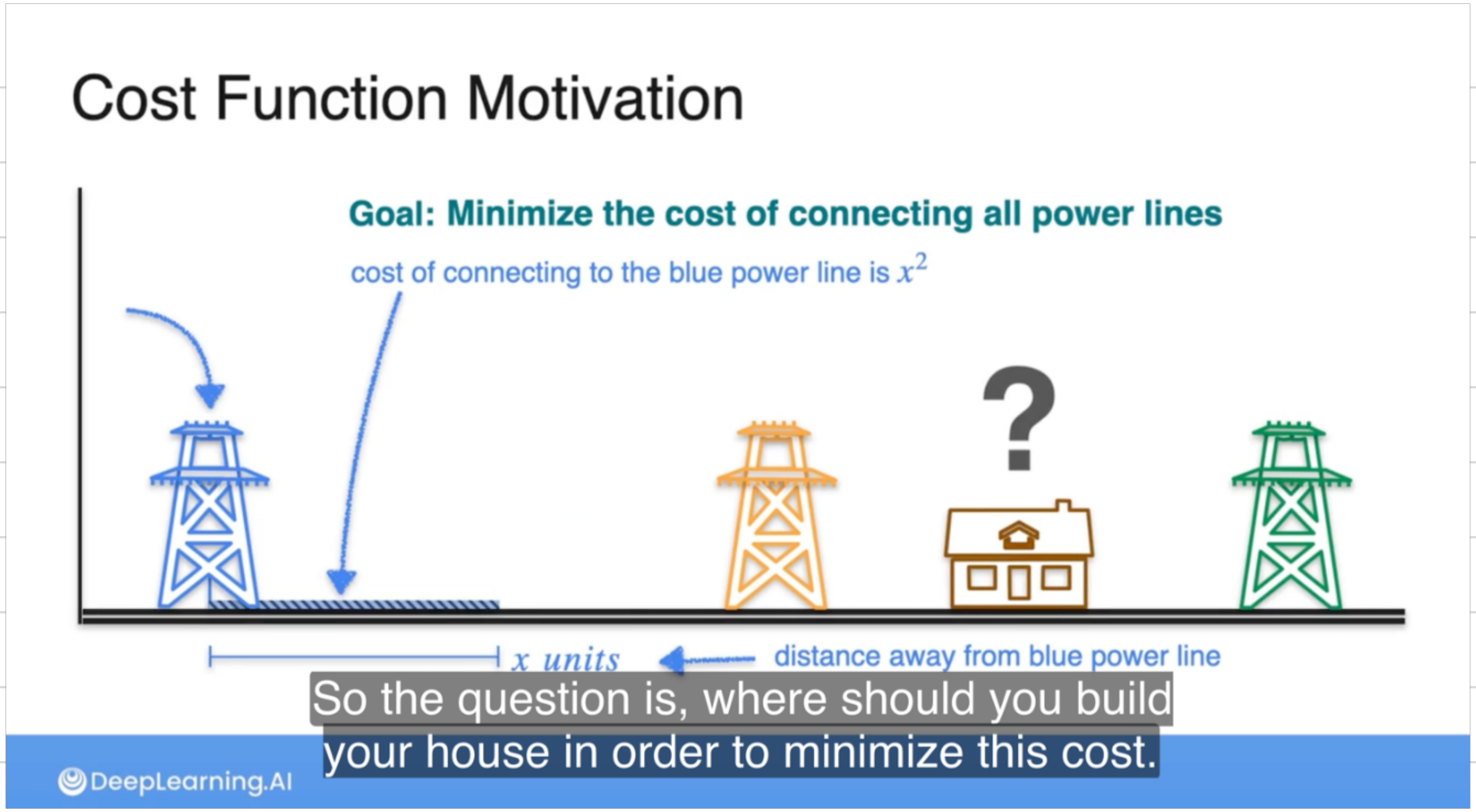
→ We could have Multiple Minima.



→ Absolute minimum → Global minimum

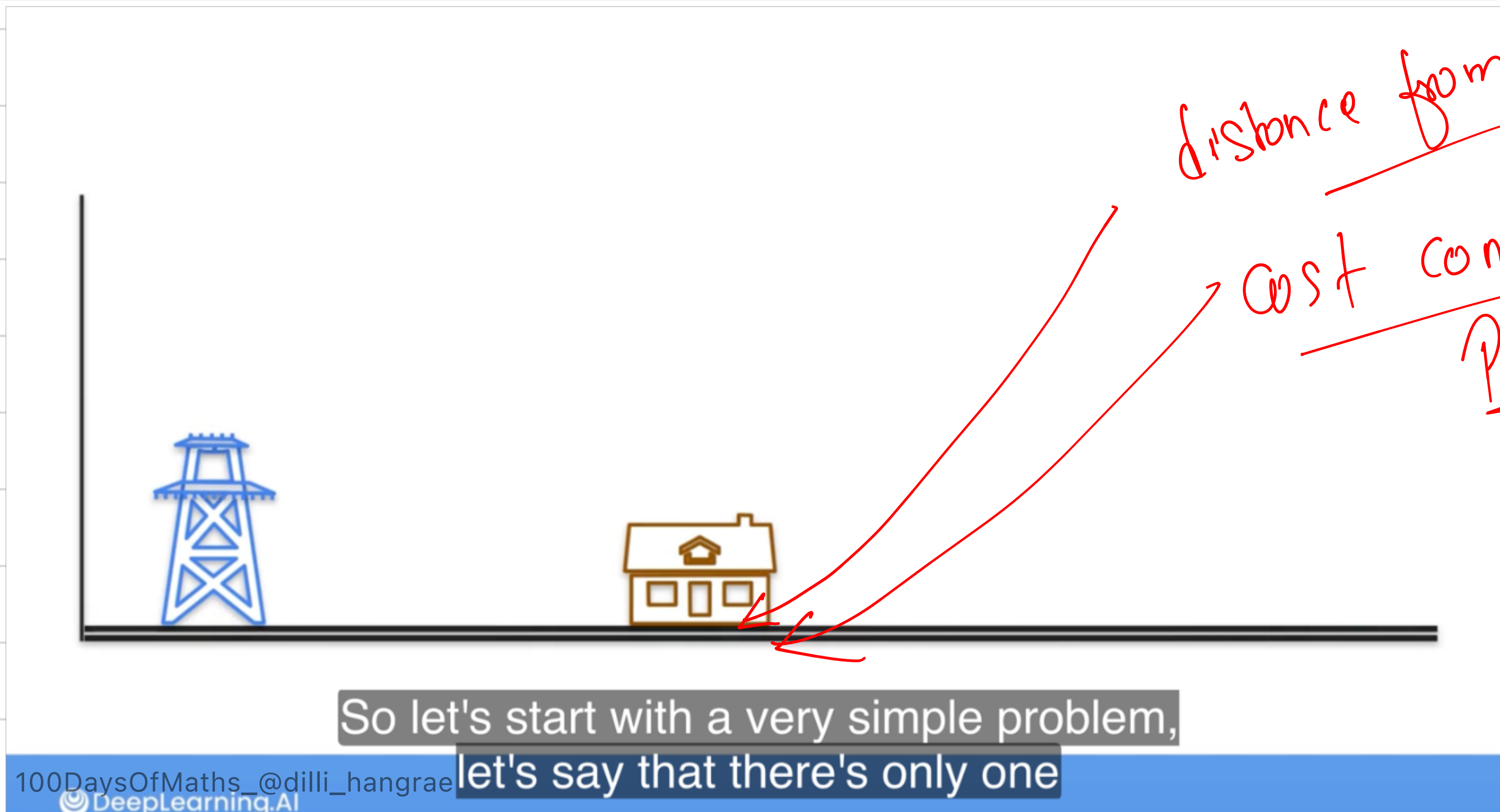
→ To optimize a function whether maximizing or minimizing it and the function is differentiable at every point, then we know one thing, it's that the candidates for max and min are those points for which derivative is zero.

Squared loss or the Squared error are very similar to each other in Machine learning minimizing the error.

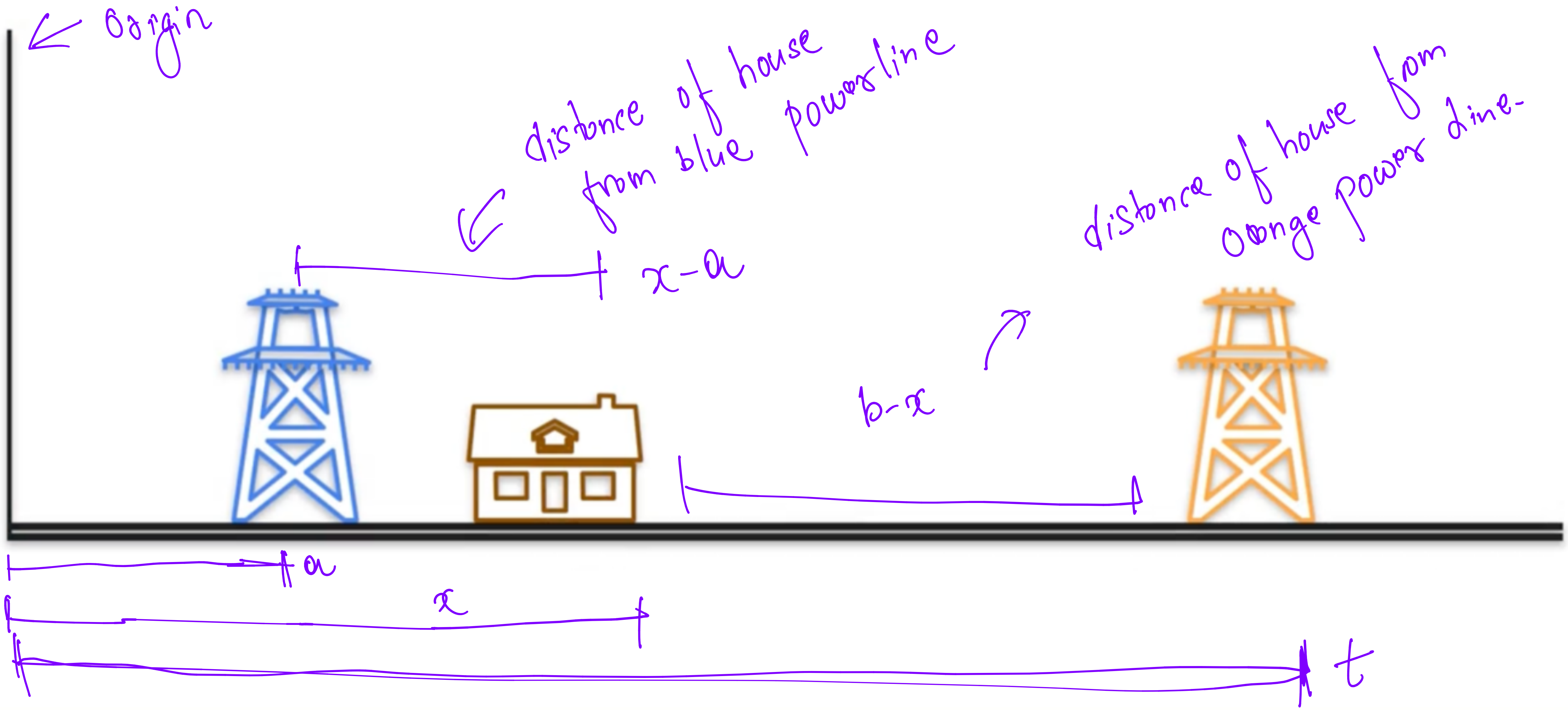


distance from blue Power line = 0 units.

cost connecting to the Power line is $0^2 = 0$



Two Power Line Problem

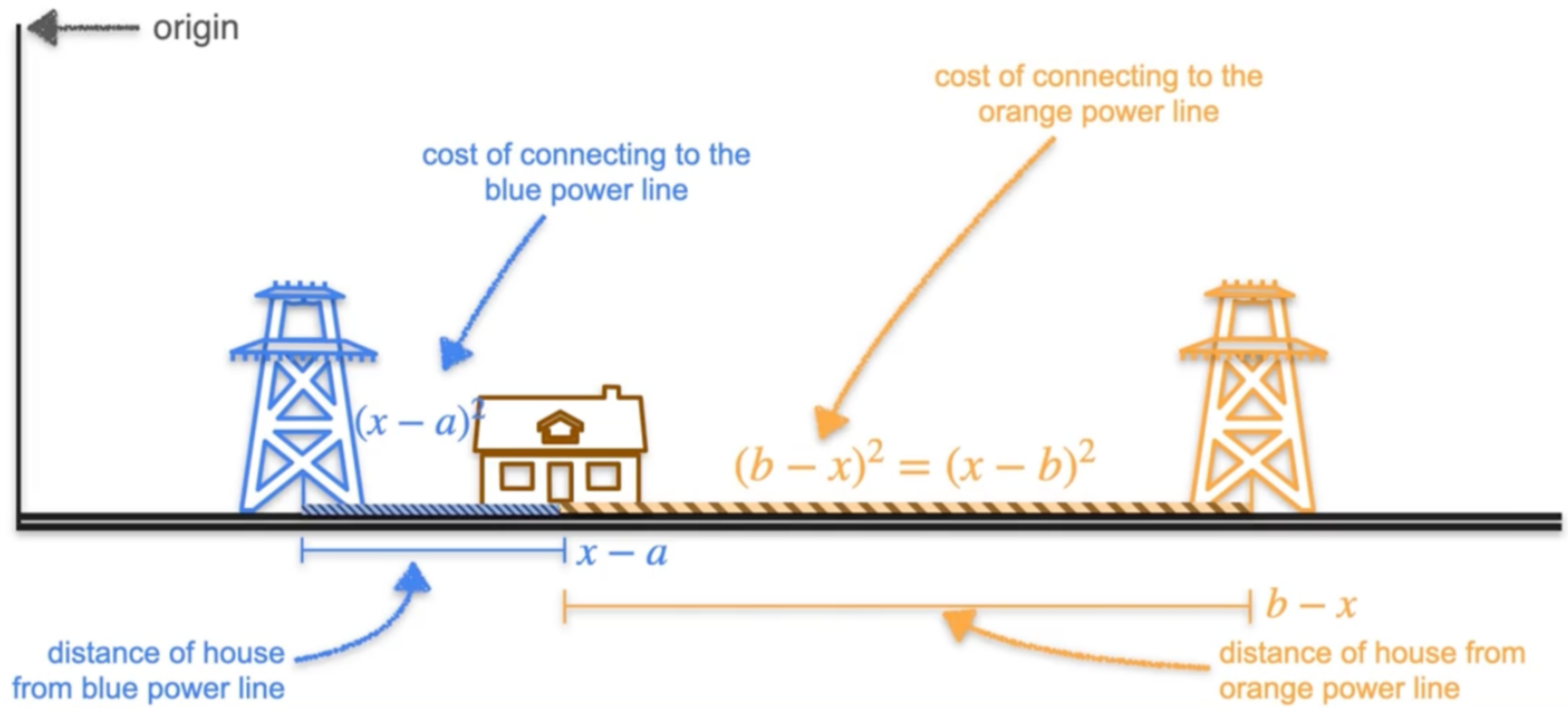


And for reference, here is the origin.

DeepLearning.AI

finding the cost of connecting the power lines between house and orange
also house and blue ?

Two Power Line Problem



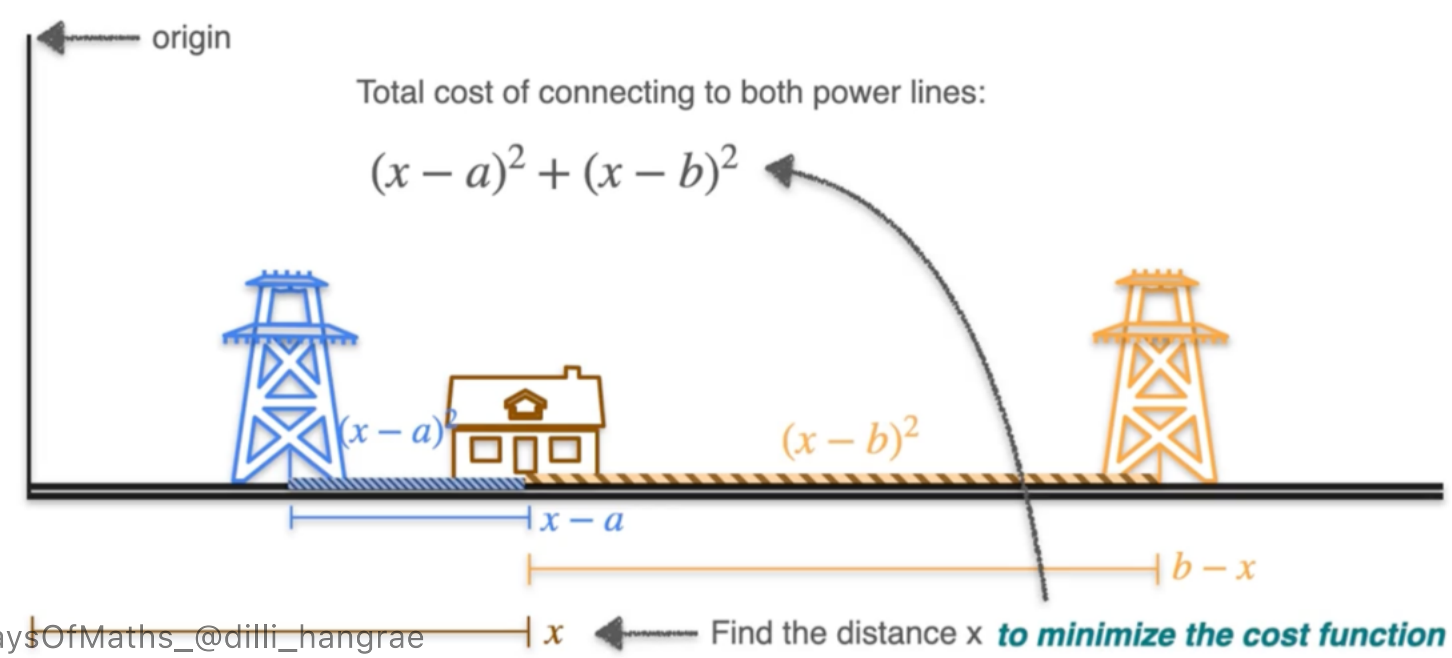
Total cost of connecting to both power lines:

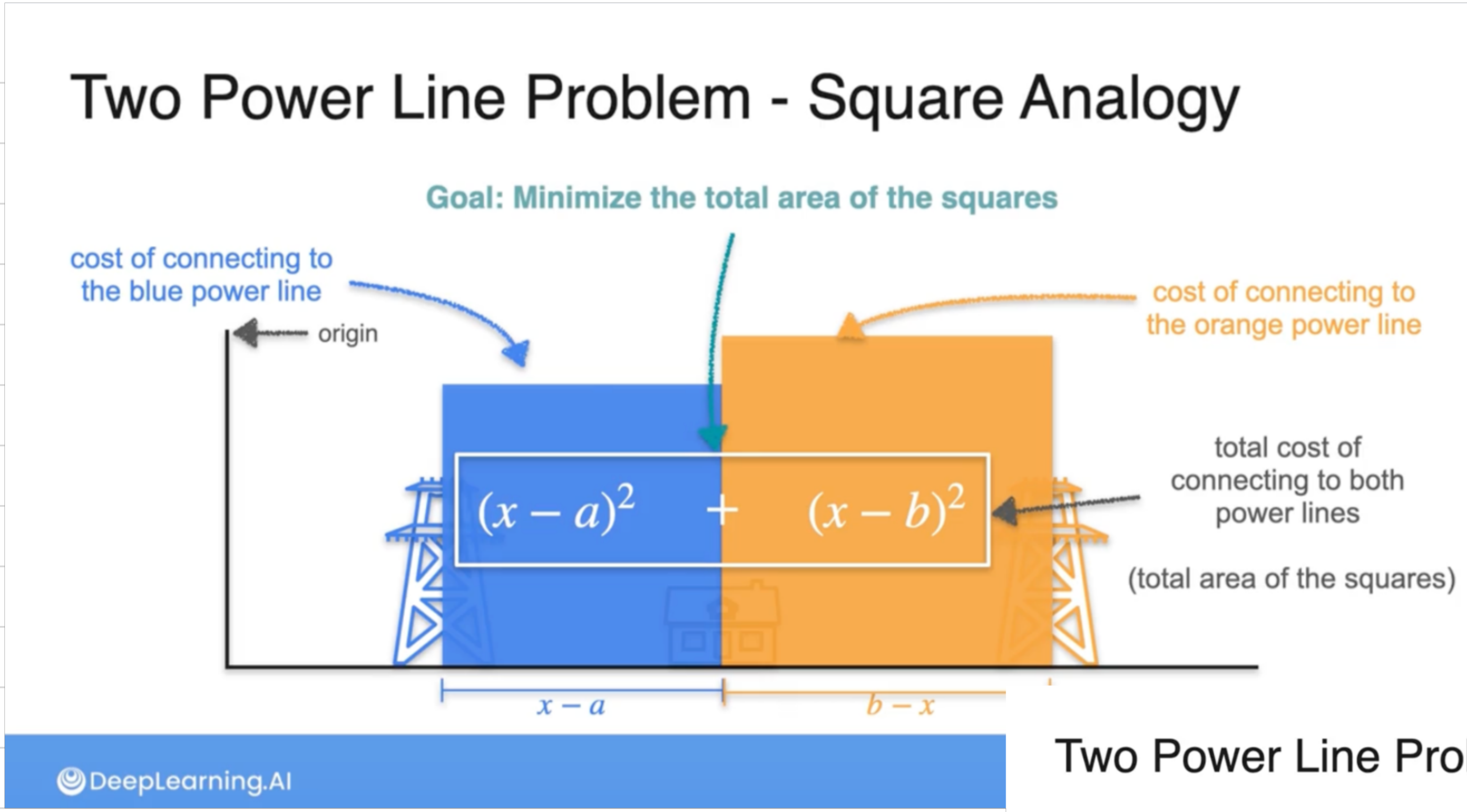
$$(x-a)^2 + (x-b)^2$$

Solution:

Obviously, the Solution is to put the House at the middle of two powers that minimize or reduce the cost.

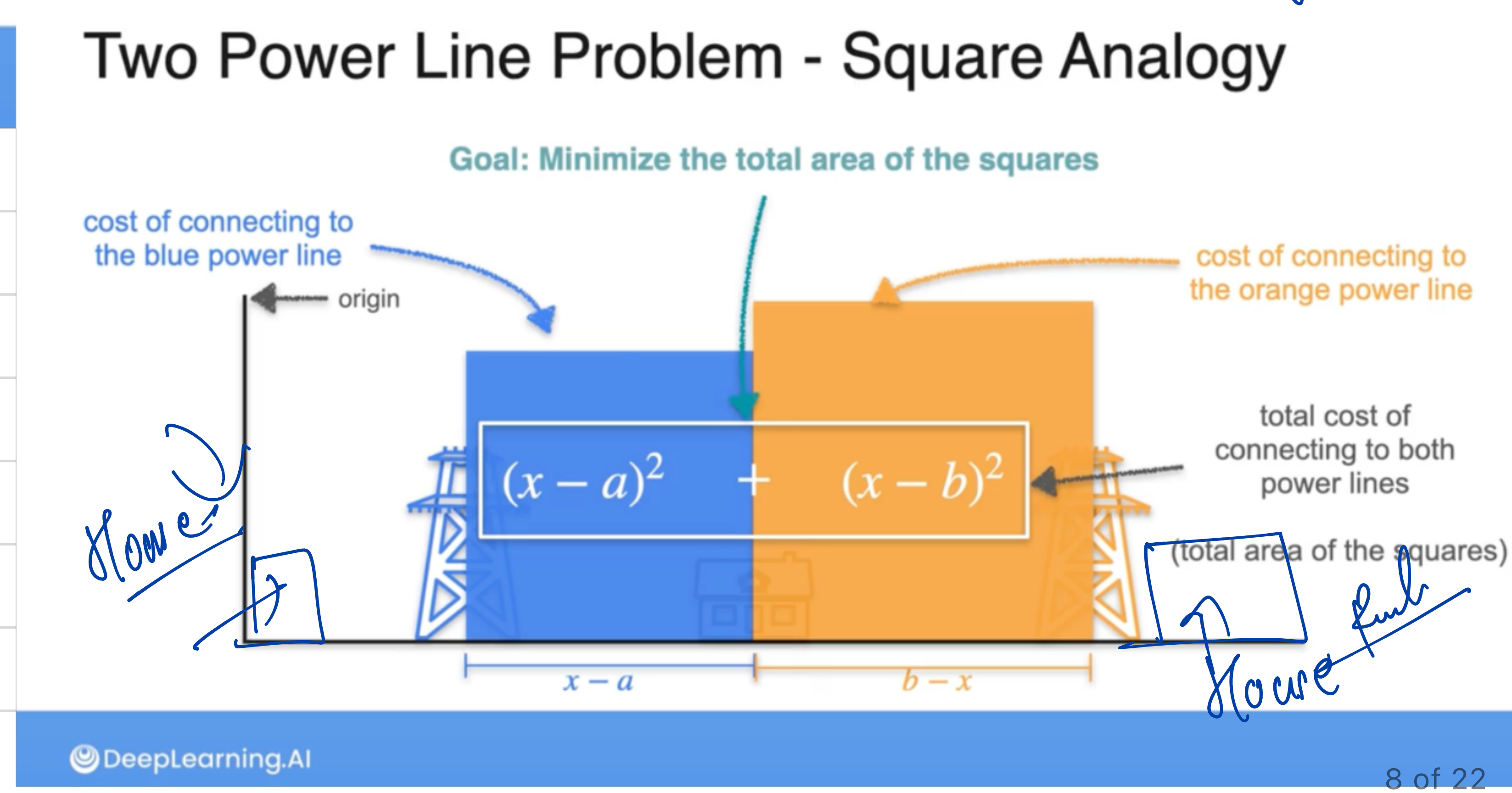
Two Power Line Problem





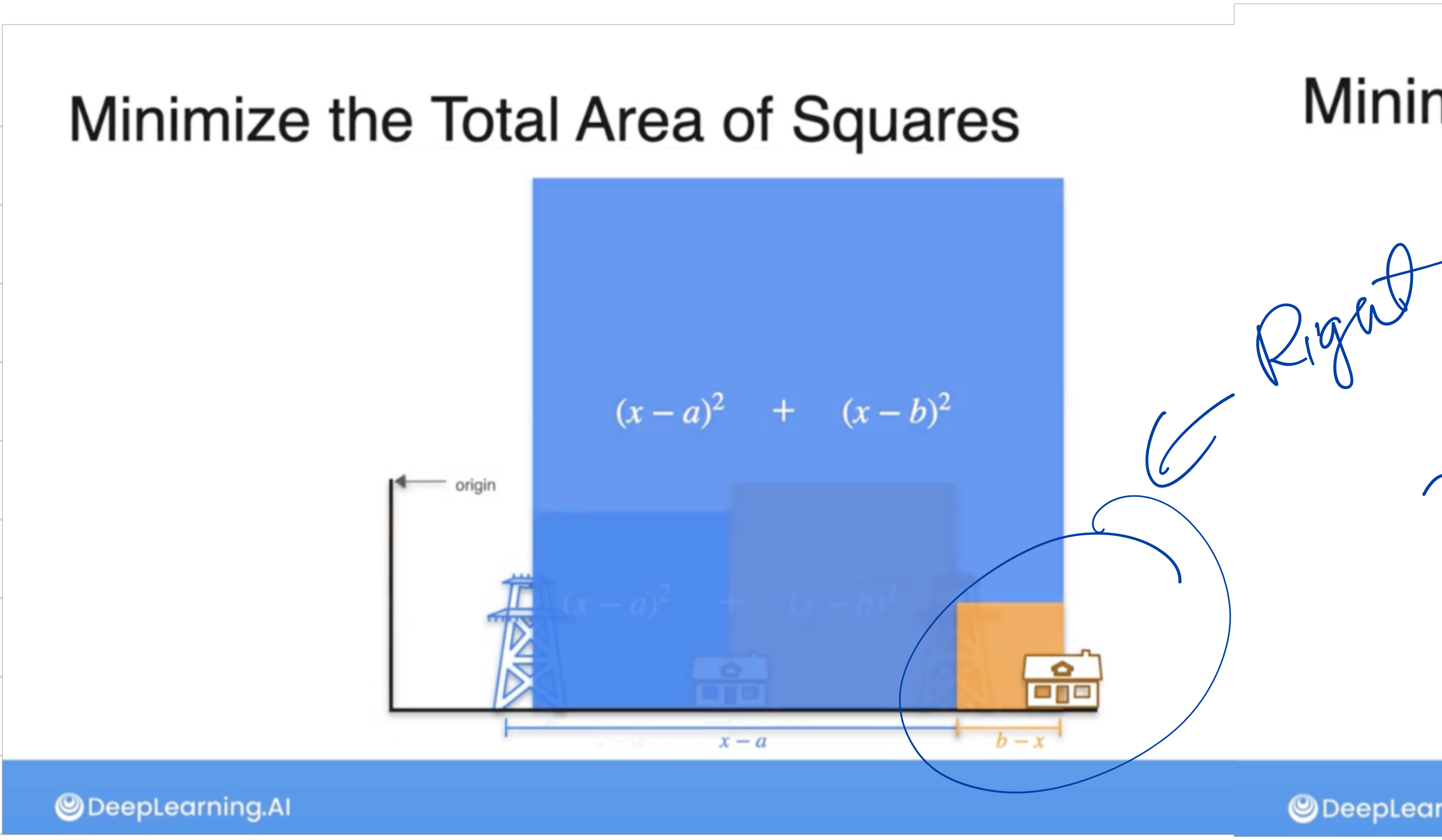
→ Affects the Total Area of Squares : If shifting the house at the left side or of the Right side.

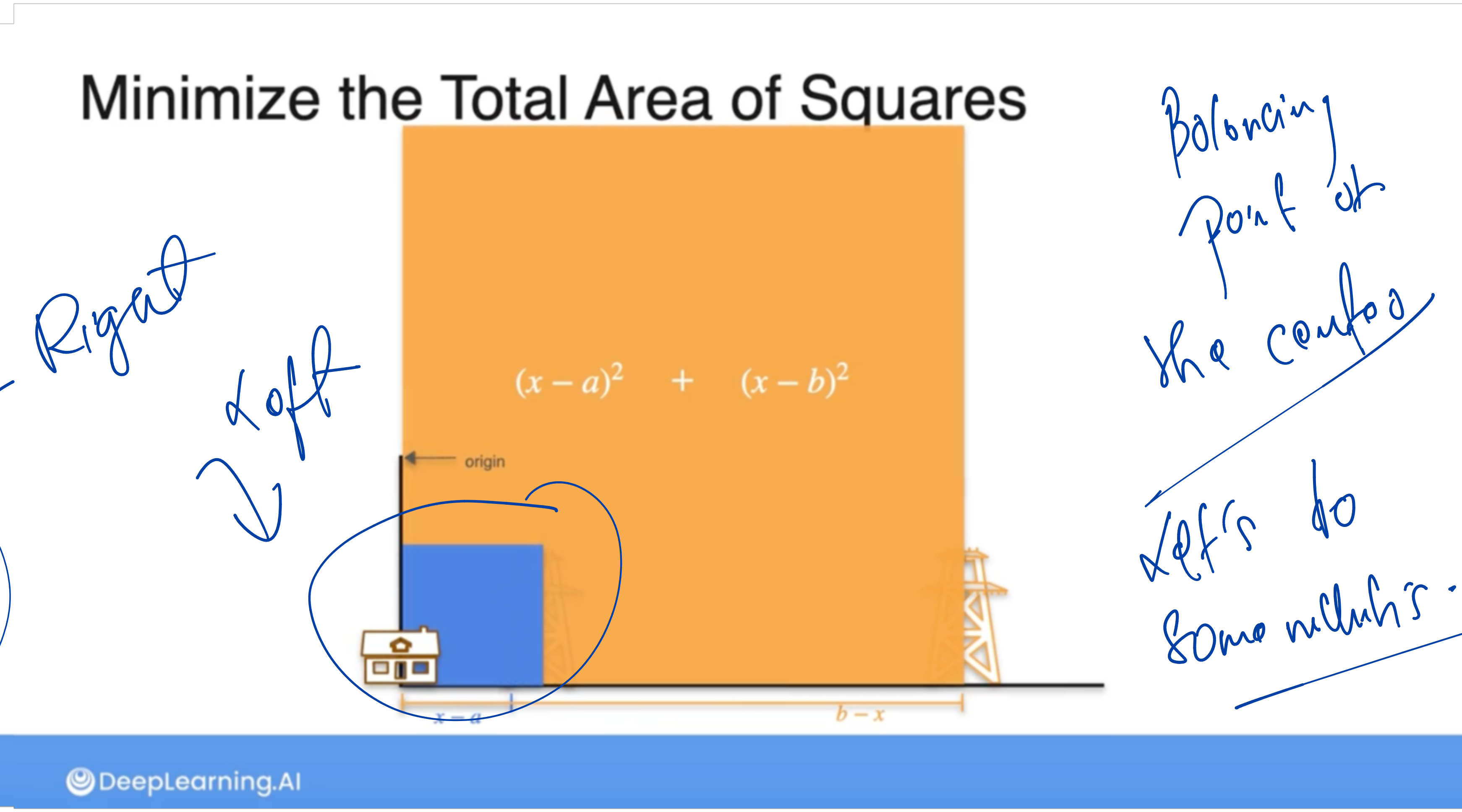
(House - either of left and
House either at Right)



House

House final





$$(x-a)^2 + (x-b)^2$$

$$\frac{d}{dx}\left[(x-a)^2 + (x-b)^2\right] = 0$$

or, $2(x-a) + 2(x-b) = 0$

or, $(x-a) + (x-b) = 0$

or, $2x - a - b = 0$

or, $2x = a + b$

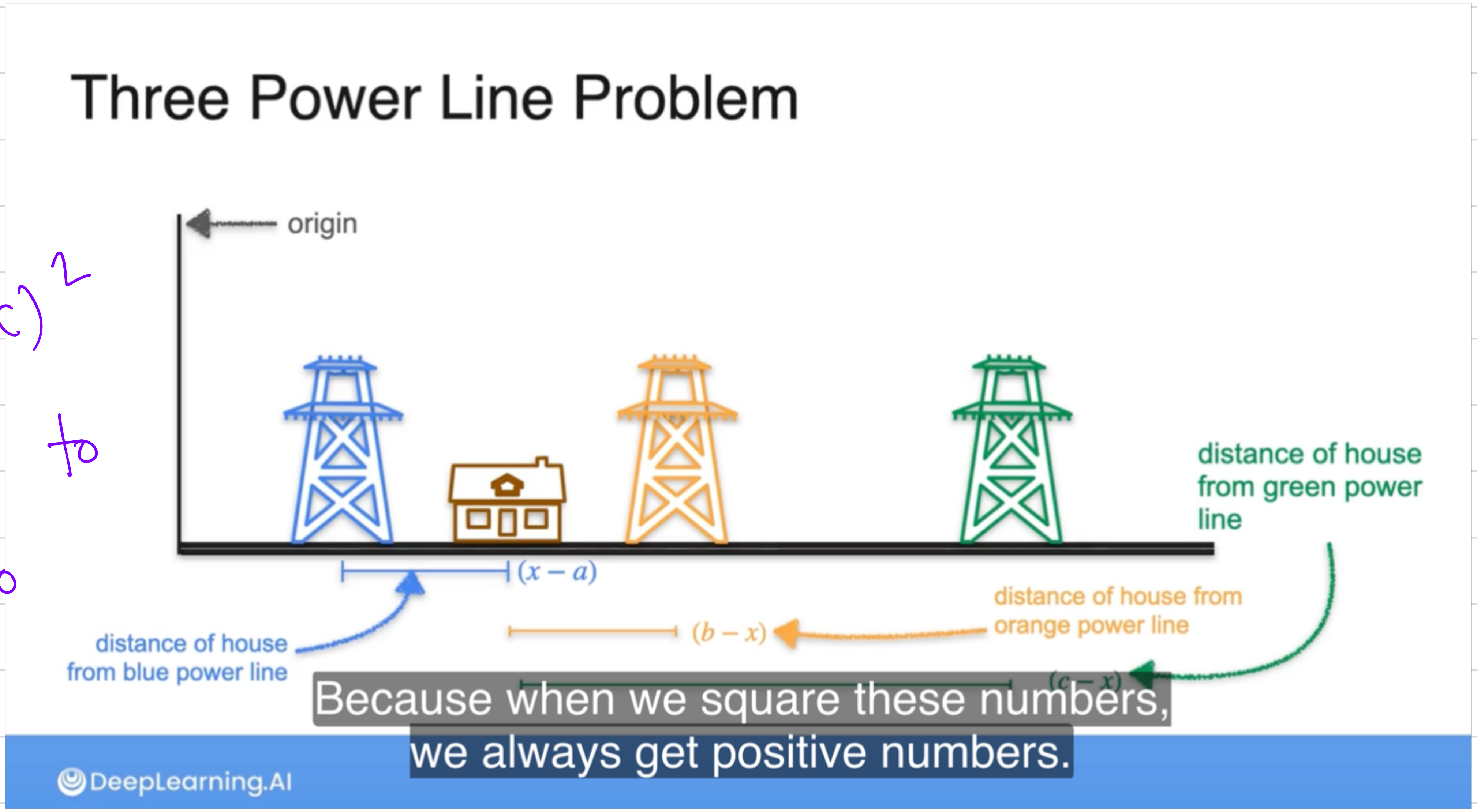
or, $x = \frac{a+b}{2}$

Three Power Line Problem

So, what is the cost function of this problem?

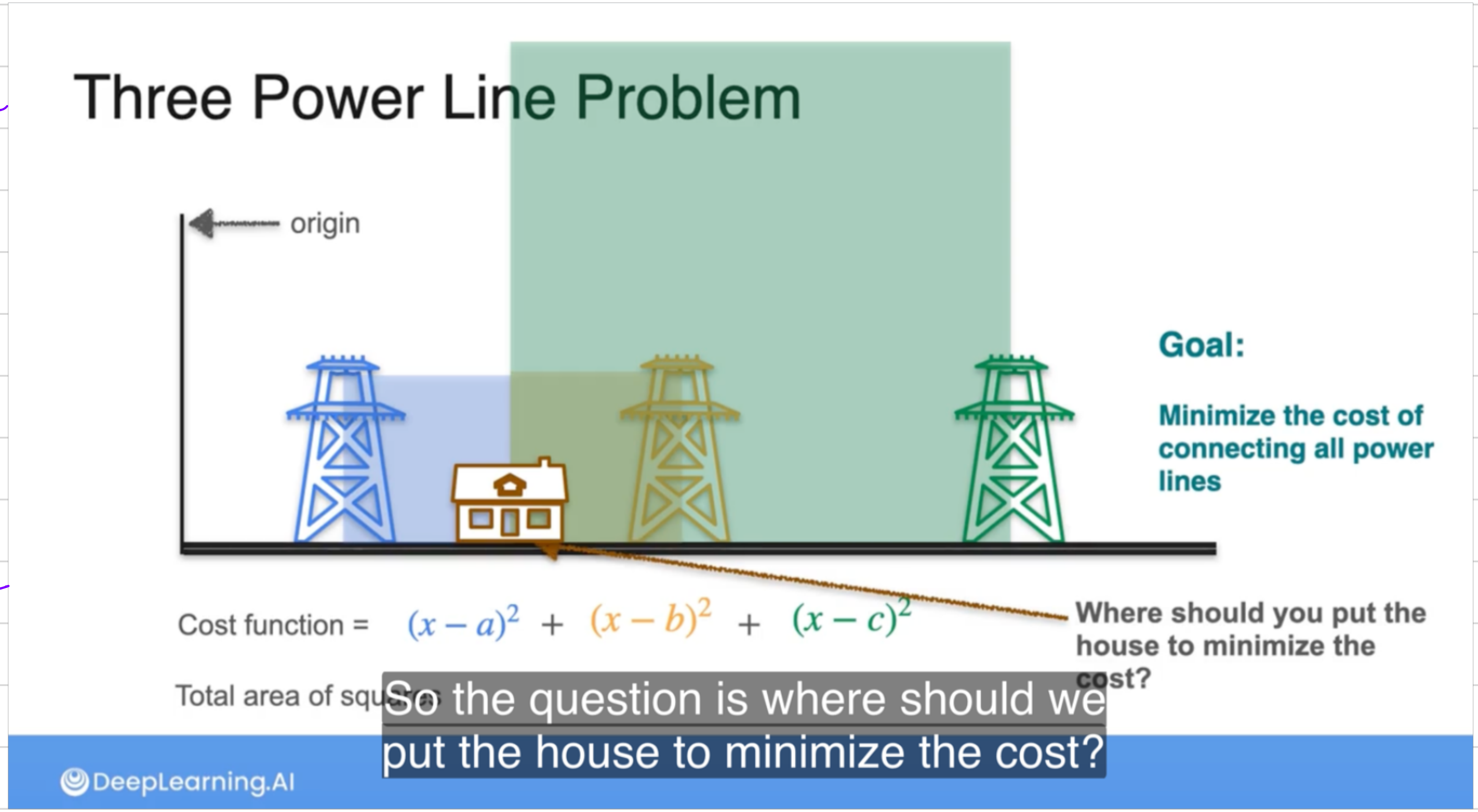
$\Rightarrow (x-a)^2 + (x-b)^2 + (x-c)^2$

$\Rightarrow$ the sum of the cost to connect to each power line, Also the total sum of squares.



Where should we put the house to minimize the cost ?

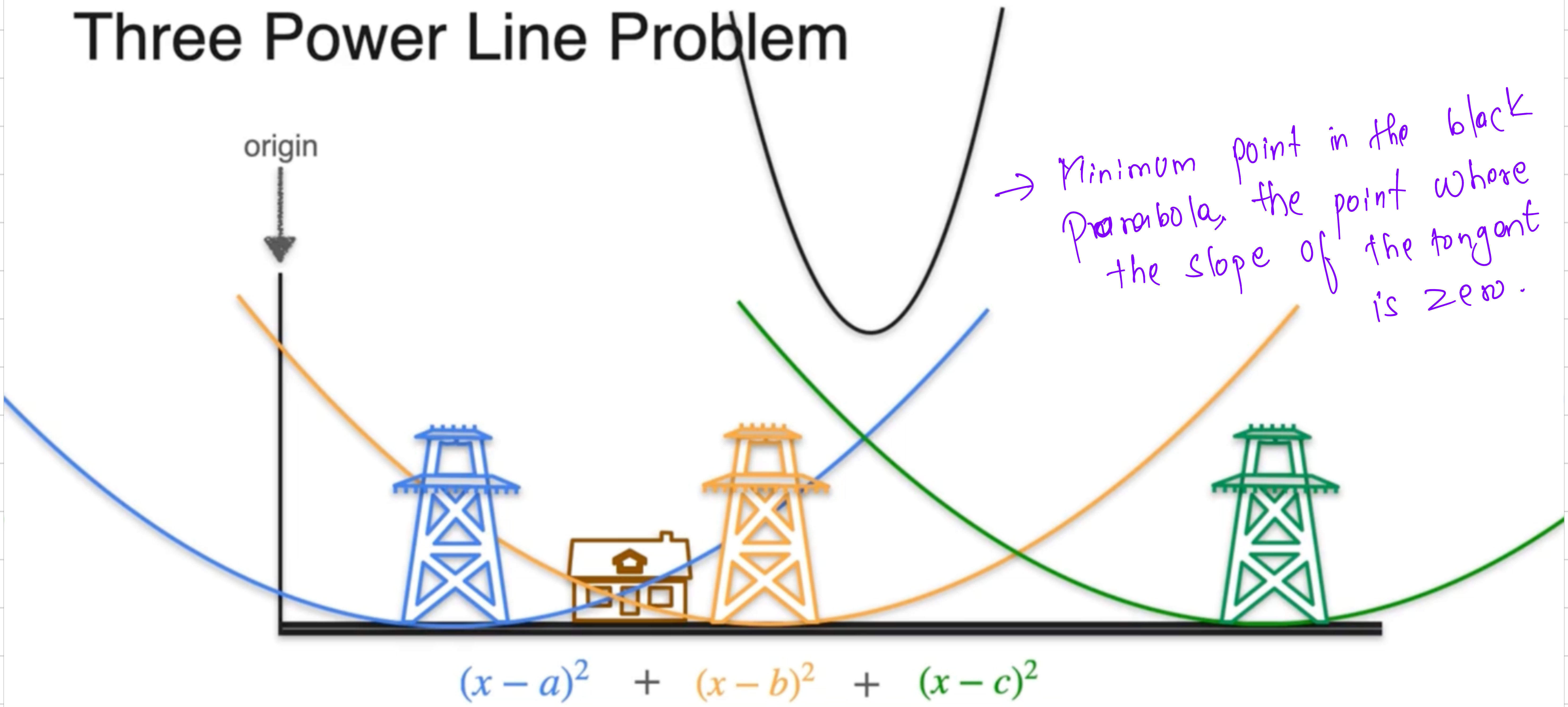
$\Rightarrow x = \frac{a+b+c}{3}$



The derivative of the cost function is $2(x-a) + 2(x-b) + 2(x-c)$

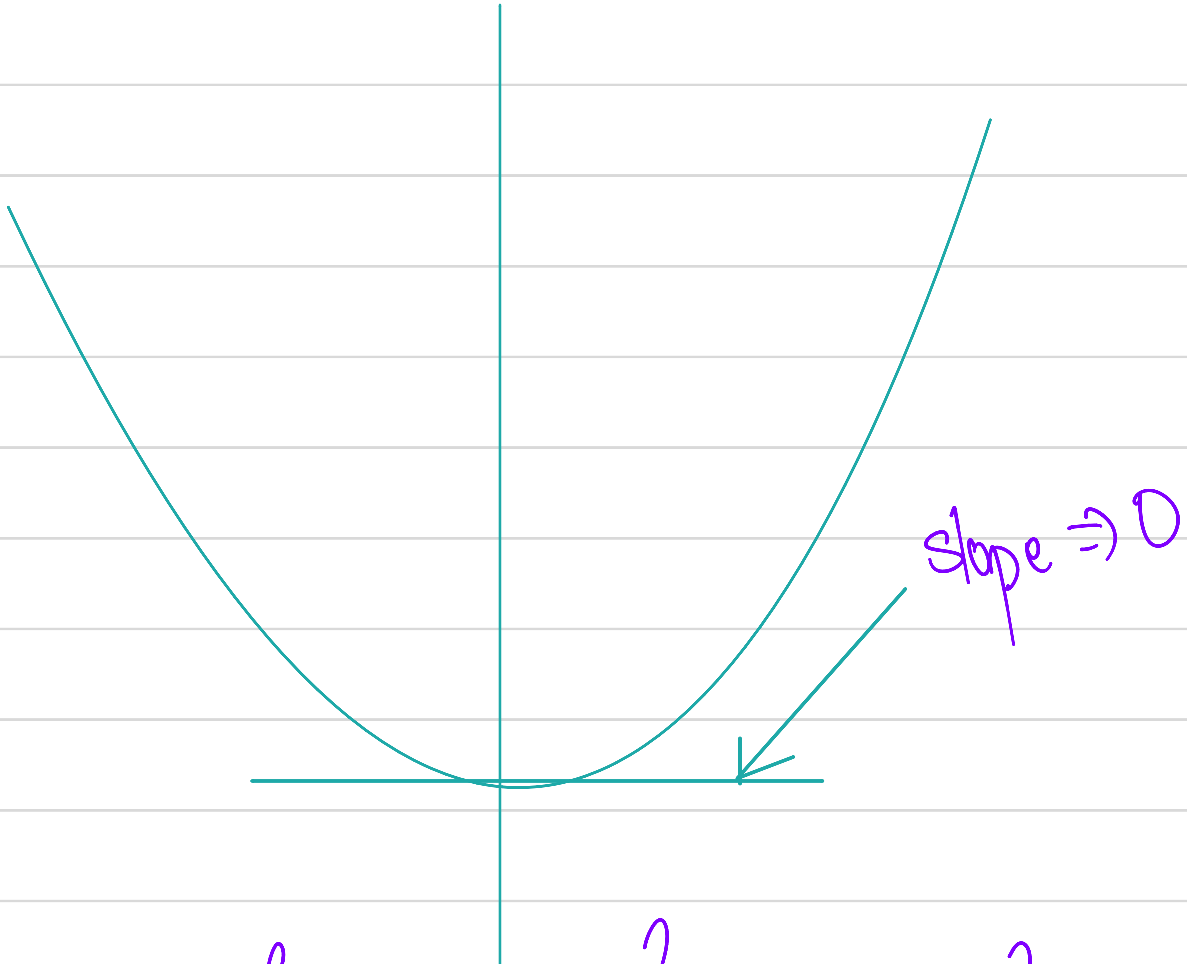
and equating this to 0 gives $x = \frac{a+b+c}{3}$.

Three Power Line Problem



→ Minimum point in the black Parabola, the point where the slope of the tangent is zero.

DeepLearning.AI



$$(x-a)^2 + (x-b)^2 + (x-c)^2$$

$$\frac{d}{dx}\left[(x-a)^2 + (x-b)^2 + (x-c)^2\right] = 0$$

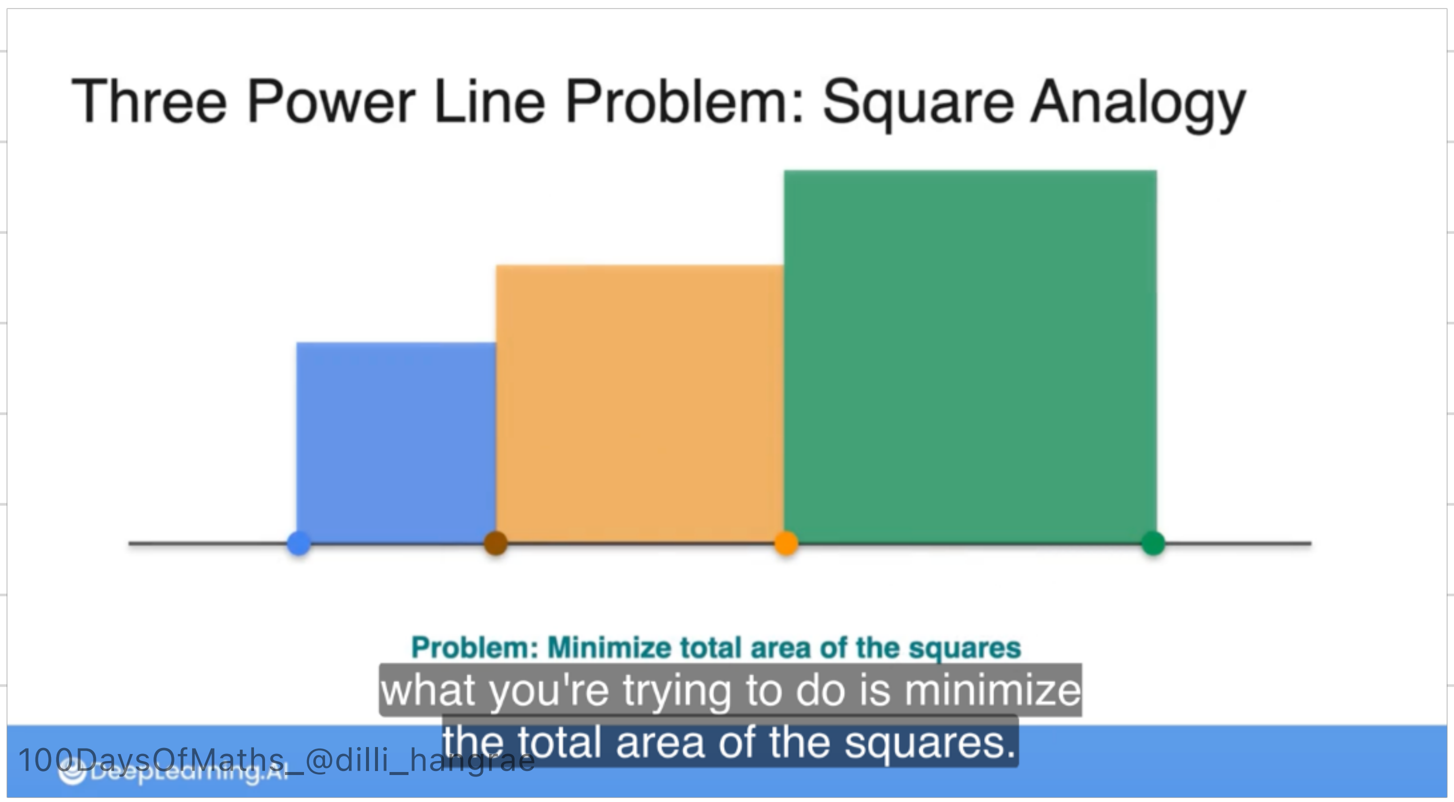
$$2(x-a) + 2(x-b) + 2(x-c) = 0$$

$$(x-a) + (x-b) + (x-c) = 0$$

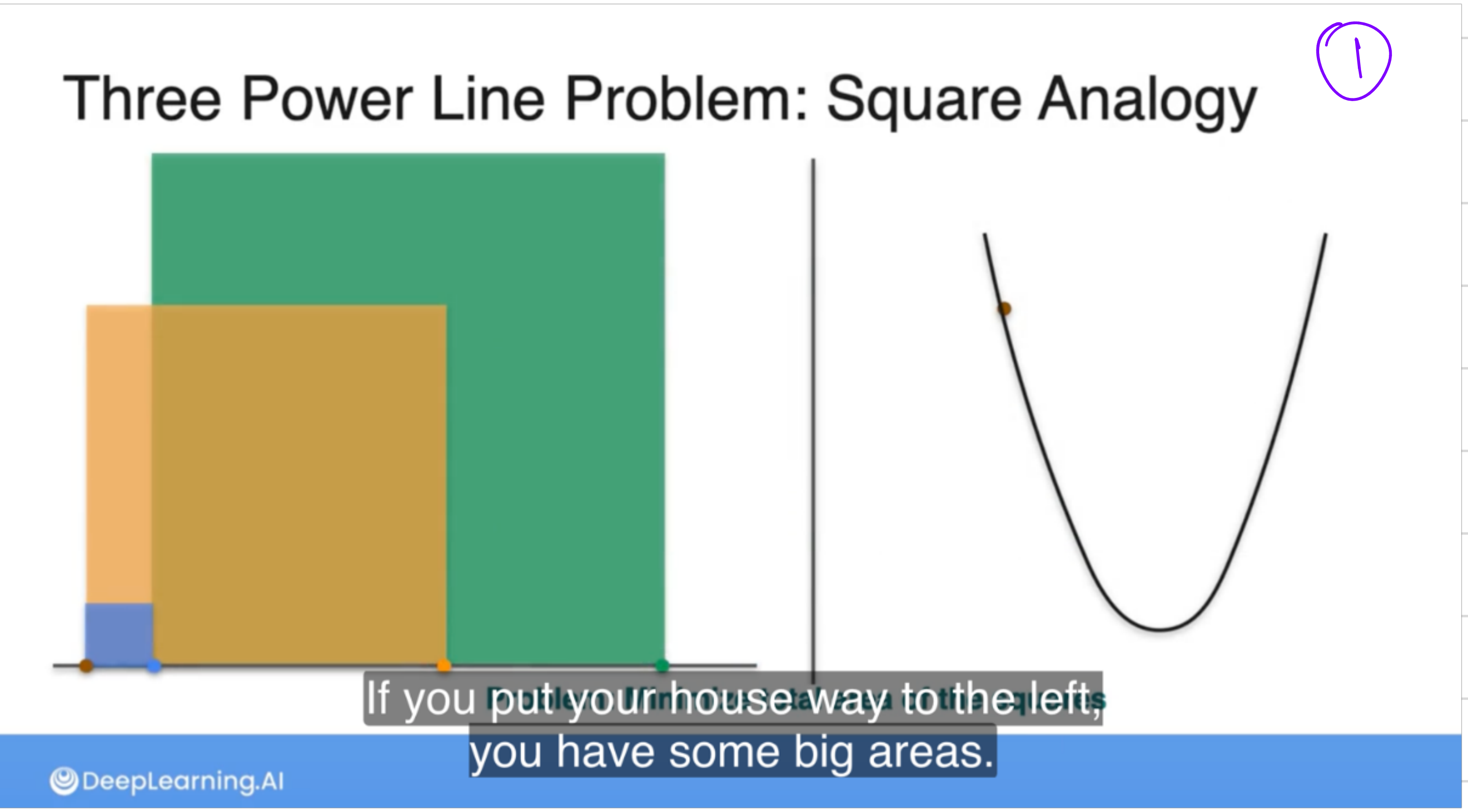
$$3x - a - b - c = 0$$

$$3x = a+b+c$$

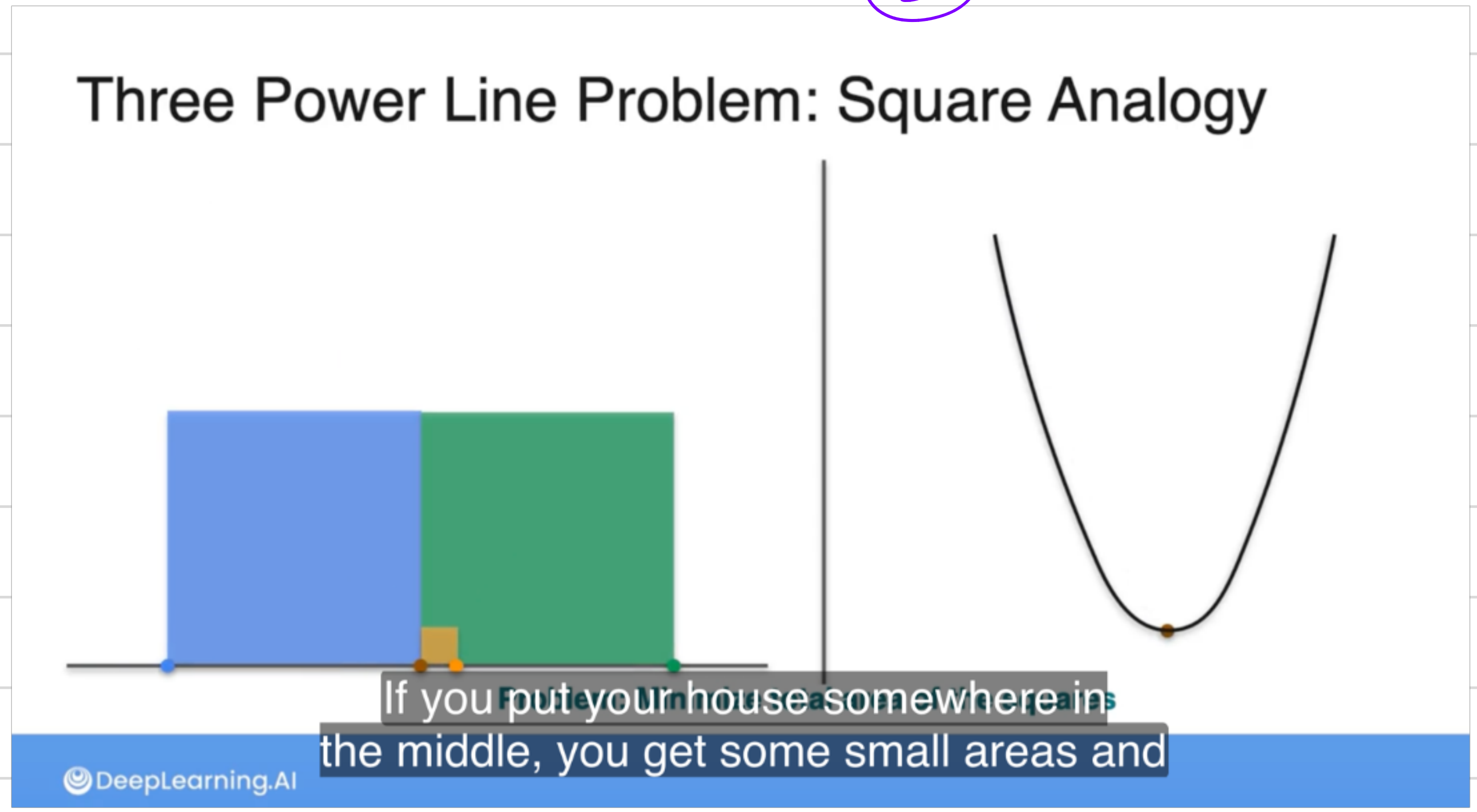
$$\boxed{x = \frac{a+b+c}{3}}$$



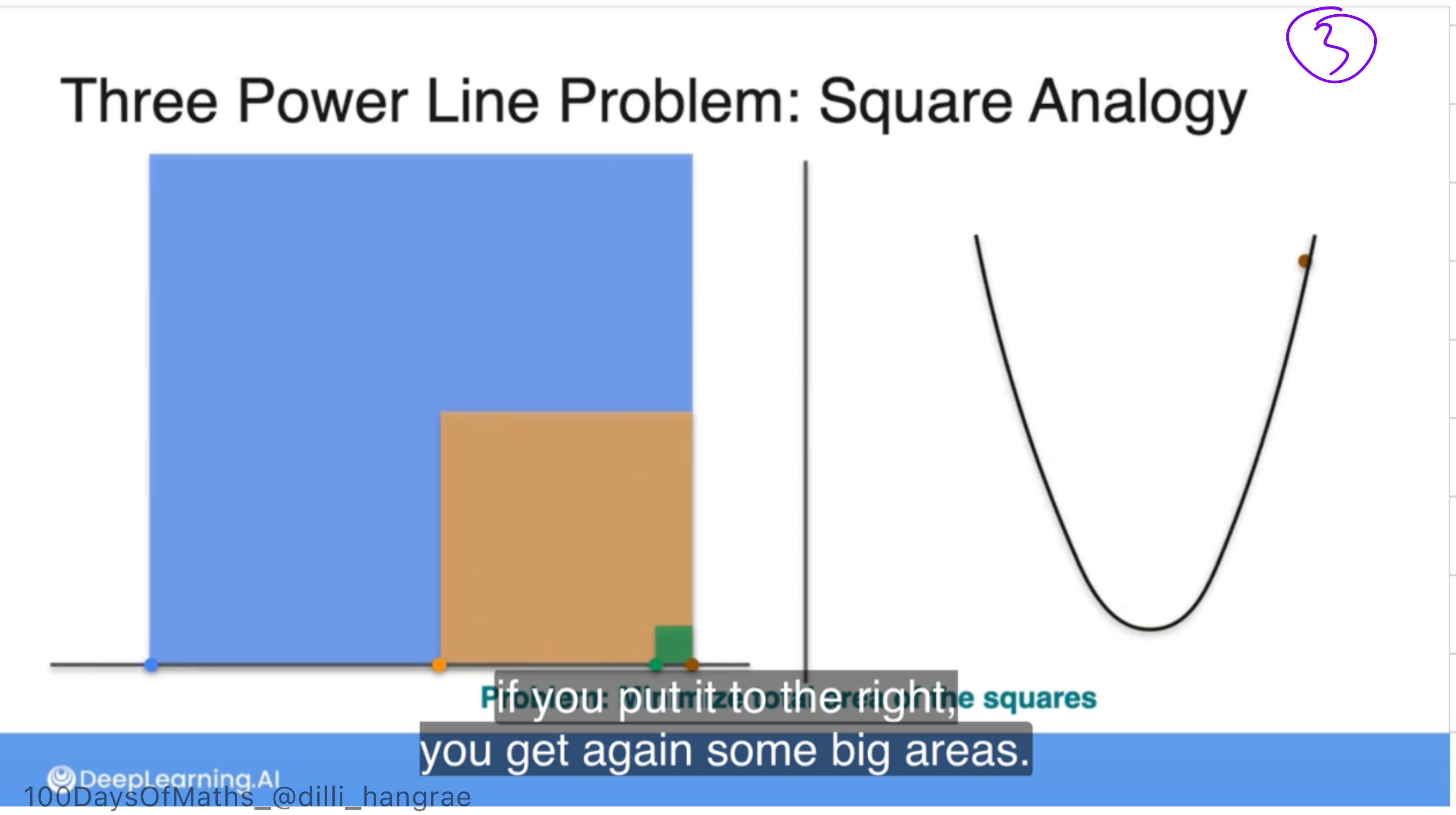
1



2



3



3' $\rightarrow$ Conditions.

Minimize $(x-a_1)^2 + (x-a_2)^2 + \dots + (x-a_n)^2$

Solution:

$$\Rightarrow x = \frac{a_1 + a_2 + \dots + a_n}{n}$$

Coin Toss

Coin 1 Coin 2 Coin 3

70% 30% 50% 50% 30% 70%

This one has a probability of landing in heads

DeepLearning.AI

Quiz

- Which of the three coins would you choose to maximize your chances of winning?

Coin 1 Coin 2 Coin 3

70% 30% 50% 50% 30% 70%

7H 3T

choose to maximize your chances of winning?

DeepLearning.AI

Because of independence we multiply

$0.7 \times 0.7 \times 0.7 \times 0.7 \times 0.7 \ldots \Rightarrow 0.7^7 \times 0.3^3$

$\Rightarrow 0.00222$

Coin 2.

$0.5 \times 0.5 \times \ldots \rightarrow 0.5^7 \, 0.5^3$

$\Rightarrow 0.00097$

Coin 3 $\rightarrow 0.3 \times 0.3 \times \ldots \Rightarrow 0.3^7 \times 0.7^3$

$\Rightarrow 0.00008$

Coins Toss

Coin 1 $0.7 \times 0.7 \times 0.7 \times 0.7 \times 0.7 \times 0.7 \times 0.7 \times 0.3 \times 0.3 \times 0.3 = 0.7^7 0.3^3$

$= 0.00222$

That's quite a small probability,

DeepLearning.AI

100DaysOfMaths_@dilli_hangrae

Coin Toss.

H — p T — $(1-p)$

Chances of winning: $p^7(1-p)^3 \Rightarrow g(p)$

Goal: maximize $g(p)$? → take derivative.

product Rule.

$$\frac{dg}{dp} = \frac{d}{dp}\left(p^7(1-p)^3\right) \Rightarrow \frac{d(p^7)}{dp}(1-p)^3 + p^7 \frac{d((1-p)^3)}{dp}$$

chain Rule

$$\Rightarrow 7p^6(1-p)^3 + p^7\, 3(1-p)^2(-1)$$

$$\Rightarrow p^6(1-p)^2\left[7(1-p) - 3p\right]$$

$$\Rightarrow p^6 (1-p)^2 (7-10p)$$ required value is 0.

$p = \cancel{0}$ always in tails. dond

$p \neq 1$ always heads

$p = 0.7$

Easier way to do that?

Yes taking log-loss

Coin Toss.

$$\log(g(p)) = \log(p^7(1-p)^3)$$

$$\Rightarrow \log(p^7) + \log(1-p)^3$$

$$\Rightarrow 7\log(p) + 3\log(1-p)$$

$$\left[\log(g(p)) \Rightarrow G(p)\right]$$

$$\frac{dG(p)}{dp} = \frac{d}{dp}\left(7\log(p) + 3\log(1-p)\right)$$

$$\Rightarrow 7\frac{1}{p} + 3\frac{1}{1-p}(-1)$$

$$\Rightarrow \frac{7(1-p) - 3p}{p(1-p)}$$

$$\Rightarrow 0$$

where $G(p)$ is the logloss

$$7(1-p) - 3p = 0$$

$$\boxed{p = 0.7}$$

Optimization of log-loss- Part 2:

Relationship with ML

H H H H H

H H T T T

Data

H T

p $(1-p)$

Model

Minimized log-loss

$p = 0.7$

Why the logarithm?

(1) Derivative of products is hard, derivative of sums is easy

$$(p) = p^6 (1-p)^2 (3-p)^9 (p-4)^{13} (10-p)^{500}$$

$$\frac{df}{dp}$$

$$[6p^5](1-p)^2(3-p)^9(p-4)^{13}(10-p)^{500} +$$

$$p^6[2(1-p)](3-p)^9(p-4)^{13}(10-p)^{500}(-1) +$$

$$p^6(1-p)^2[9(3-p)^8](p-4)^{13}(10-p)^{500}(-1) +$$

$$p^6(1-p)^2(3-p)^9[13(p-4)^{12}](10-p)^{500} + p^6(1-p)^2(3-p)^9$$

$$(p-4)^{13}[500(10-p)^{499}](-1)$$

$$\Downarrow$$

$$\Leftarrow \frac{d}{dp}\log(f)$$

$$\Leftarrow \frac{6}{p} + \frac{2}{(1-p)}(-1) + \frac{9}{3-p}(-1) + \frac{13}{p-4} + \frac{500}{10-p}(-1)$$

Why Logarithm?

(Taking log helps to analyze complex Eqⁿs)

2. Product of lots of tiny things is tiny!

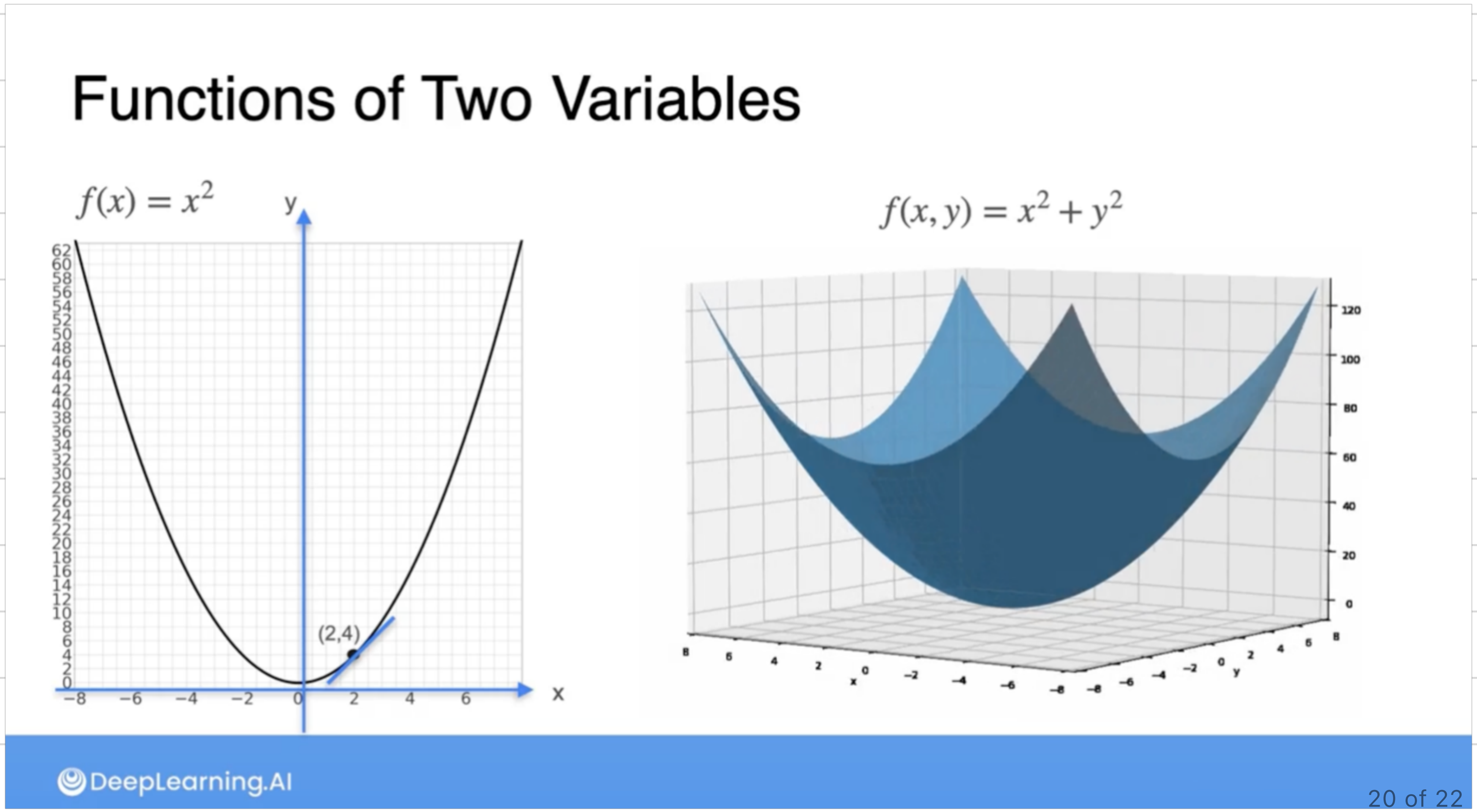
1. Derivative of products is hard, derivative of sums is easy!

Introduction to Tangent planes:

Finding the tangent plane of f(x,y) at a point.

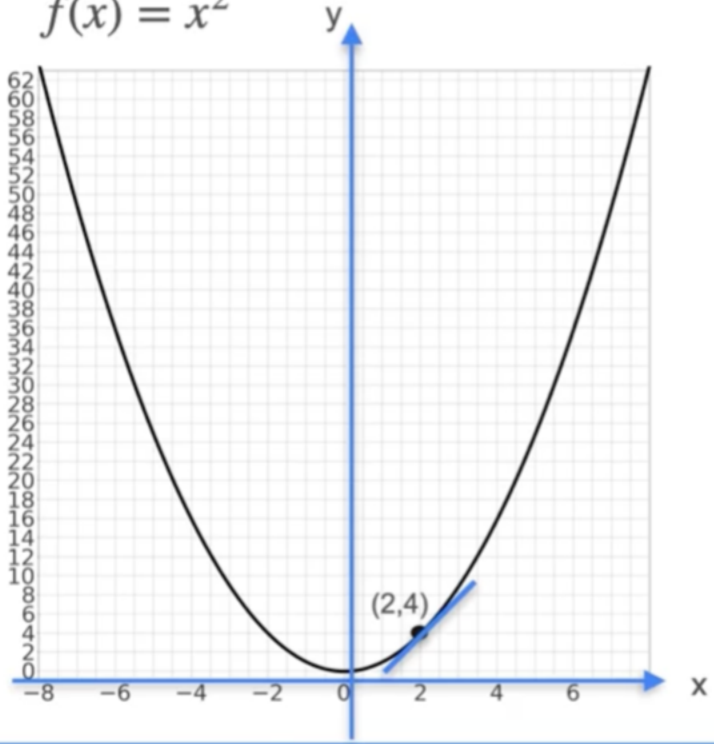
→ Given function f(x,y)

→ Point of Tangency: (a,b)



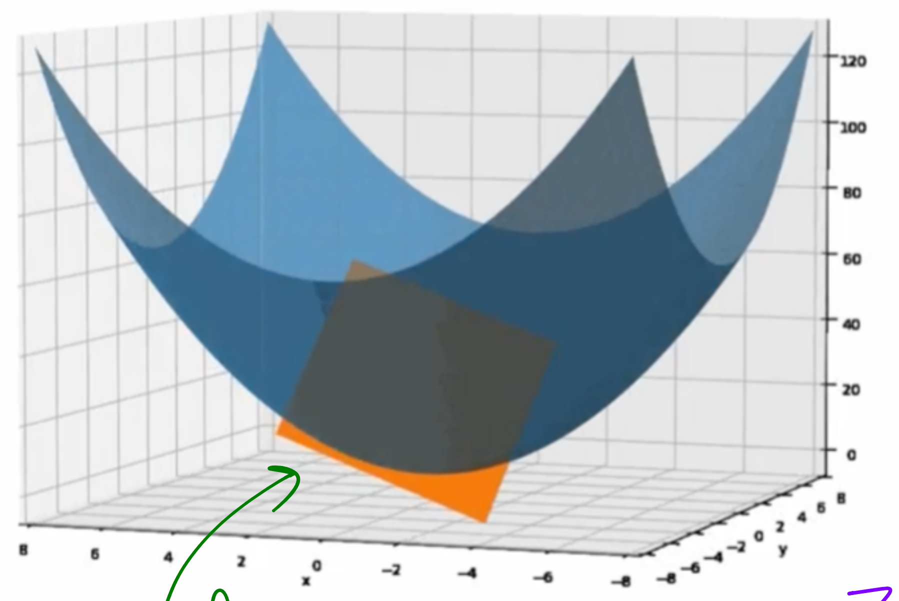
Functions of Two Variables

$f(x) = x^2$



1

$f(x, y) = x^2 + y^2$



tangent plane

DeepLearning.AI

Compute Partial Derivative

$f_x(x,y) = \frac{\partial f}{\partial x}$

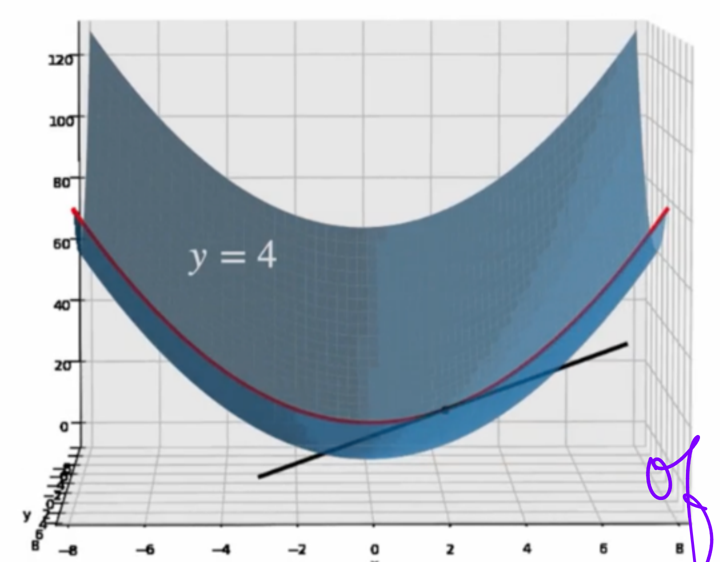
$f_y(x,y) = \frac{\partial f}{\partial y}$

Evaluate (a,b)

$f_x(a,b)$ and $f_y(a,b)$

$Z = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$

Finding the Tangent Plane



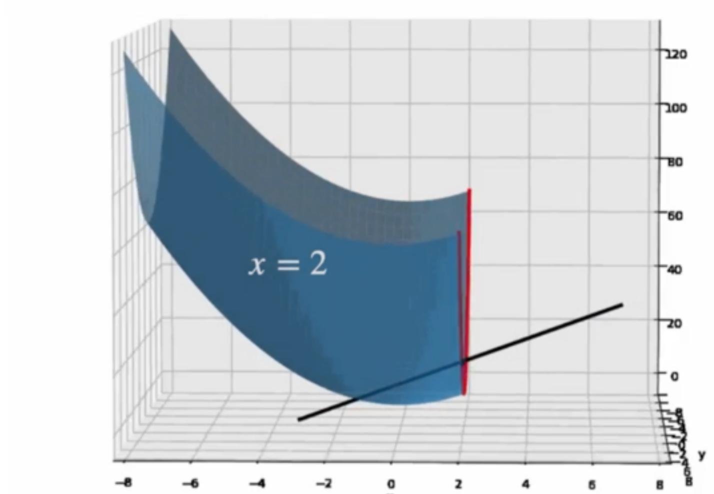
Fix y=4 $f(x,4) = x^2 + 4^2$

2

linear Approximation of $f(x,y)$ near (a,b).

DeepLearning.AI

Finding the Tangent Plane



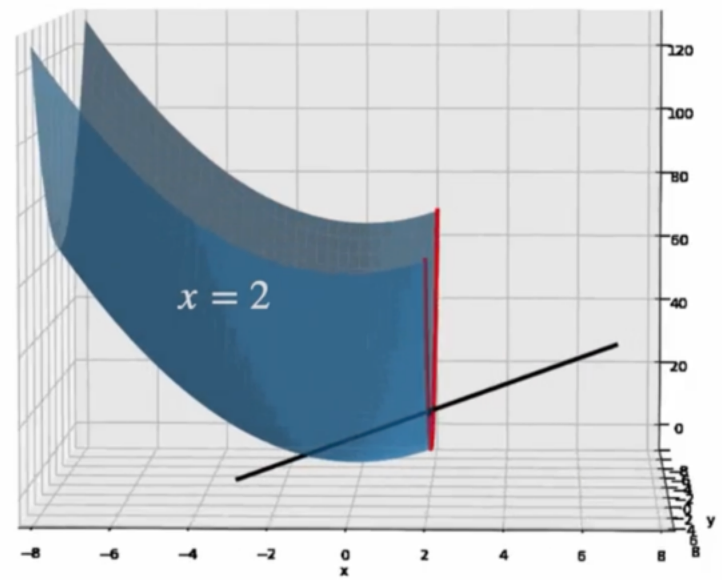
Fix y=4 $f(x,4) = x^2 + 4^2$

$$\frac{d}{dx}\big(f(x,4)\big) = 2x$$

Fix x=2 $f(2,y) = 2^2 + y^2$

3

Finding the Tangent Plane



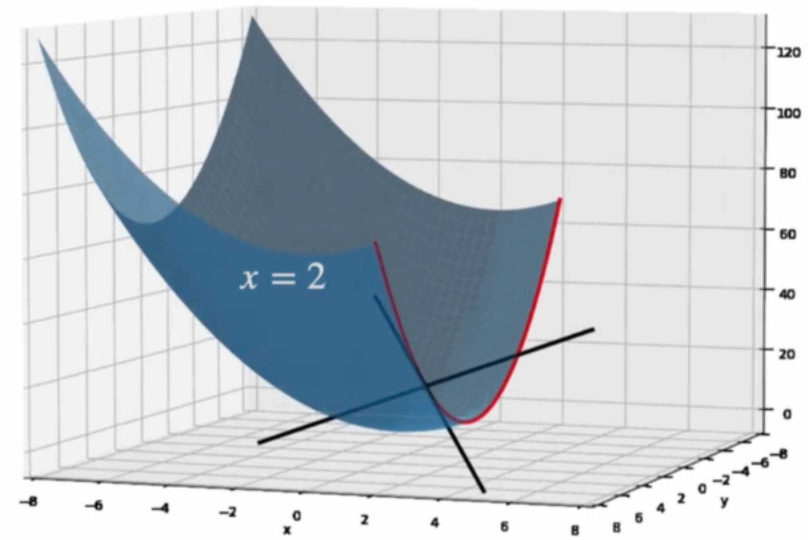
Fix y=4 $\quad f(x,4) = x^2 + 4^2$

$$\frac{d}{dx}\big(f(x,4)\big) = 2x$$

Fix x=2 $\quad f(2,y) = 2^2 + y^2$

DeepLearning.AI

Finding the Tangent Plane



Fix y=4 $\quad f(x,4) = x^2 + 4^2$

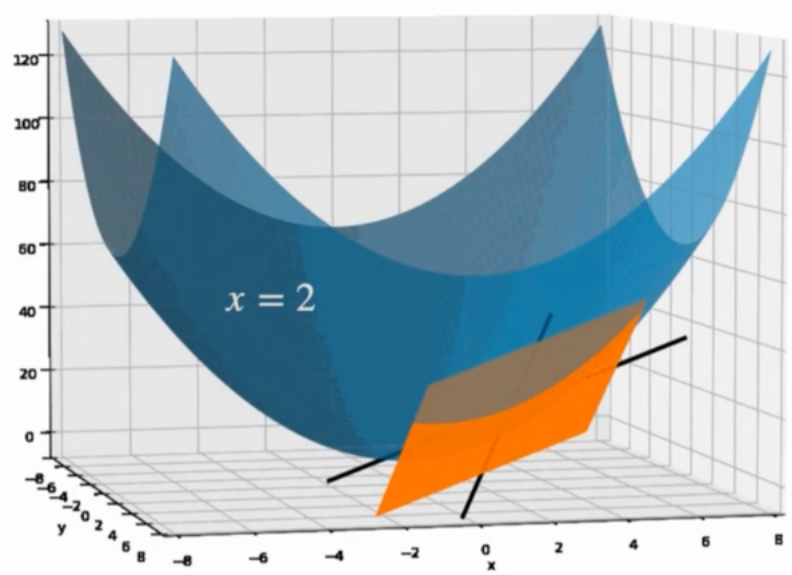
$$\frac{d}{dx}\big(f(x,4)\big) = 2x$$

Fix x=2 $\quad f(2,y) = 2^2 + y^2$

$$\frac{d}{dy}\big(f(2,y)\big) = 2y$$

DeepLearning.AI

Finding the Tangent Plane



Fix y=4 $\quad f(x,4) = x^2 + 4^2$

$$\frac{d}{dx}\big(f(x,4)\big) = 2x$$

Fix x=2 $\quad f(2,y) = 2^2 + y^2$

$$\frac{d}{dy}\big(f(2,y)\big) = 2y$$

The tangent plane contains both tangent lines.

DeepLearning.AI

Source: Calculus for Machine Learning and Data Science Coursera Course offered by DeepLearning.ai.