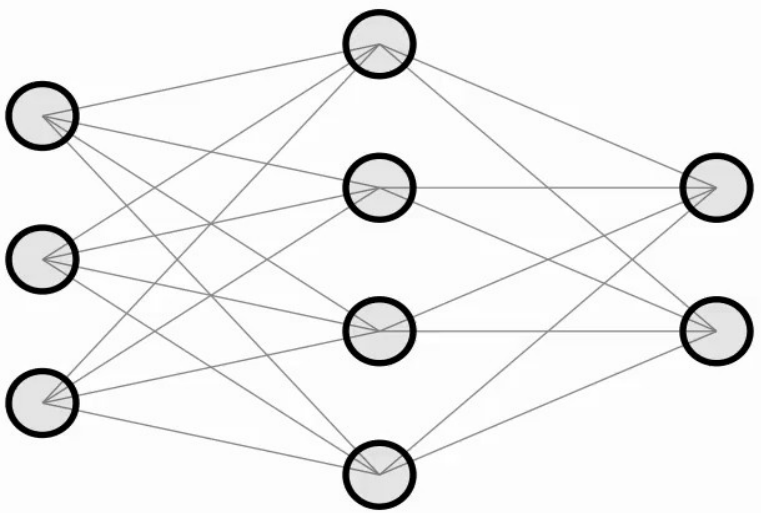
Introduction to Machine Learning

by Filip Pawlicki

What is Machine Learning?

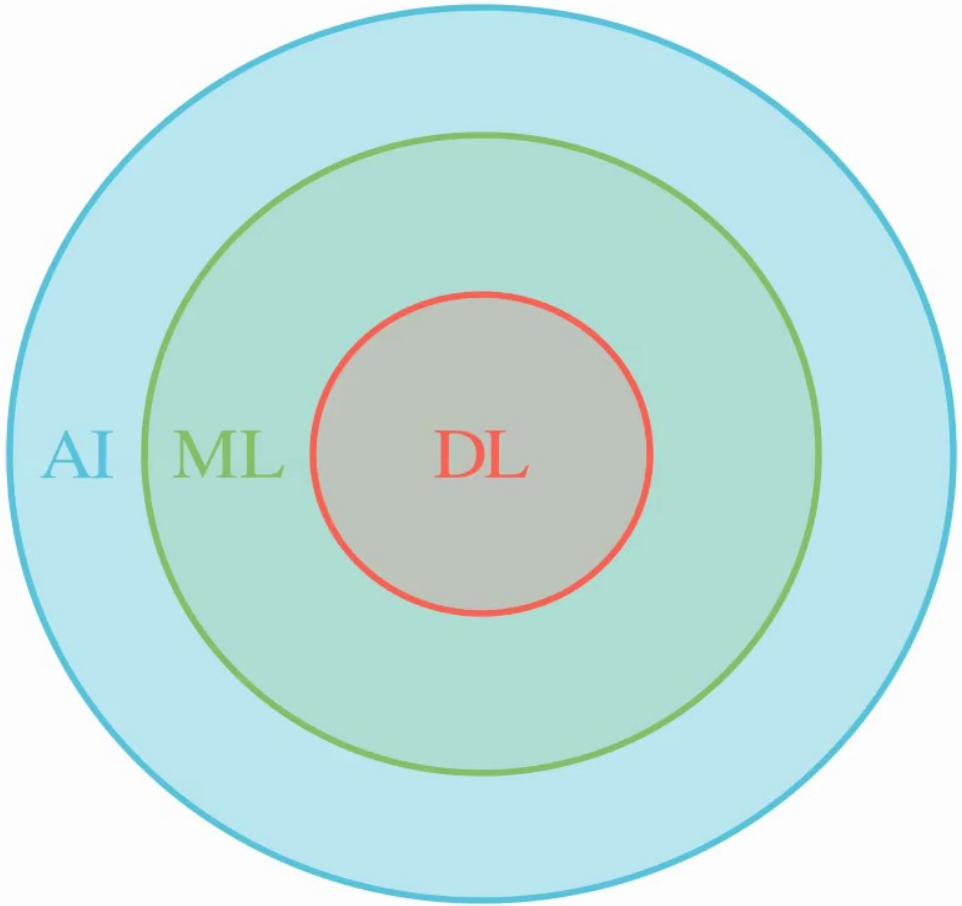
Machine Learning (ML) is a field of AI that allows computers to learn patterns from data and make predictions or decisions.

"Best science club?"

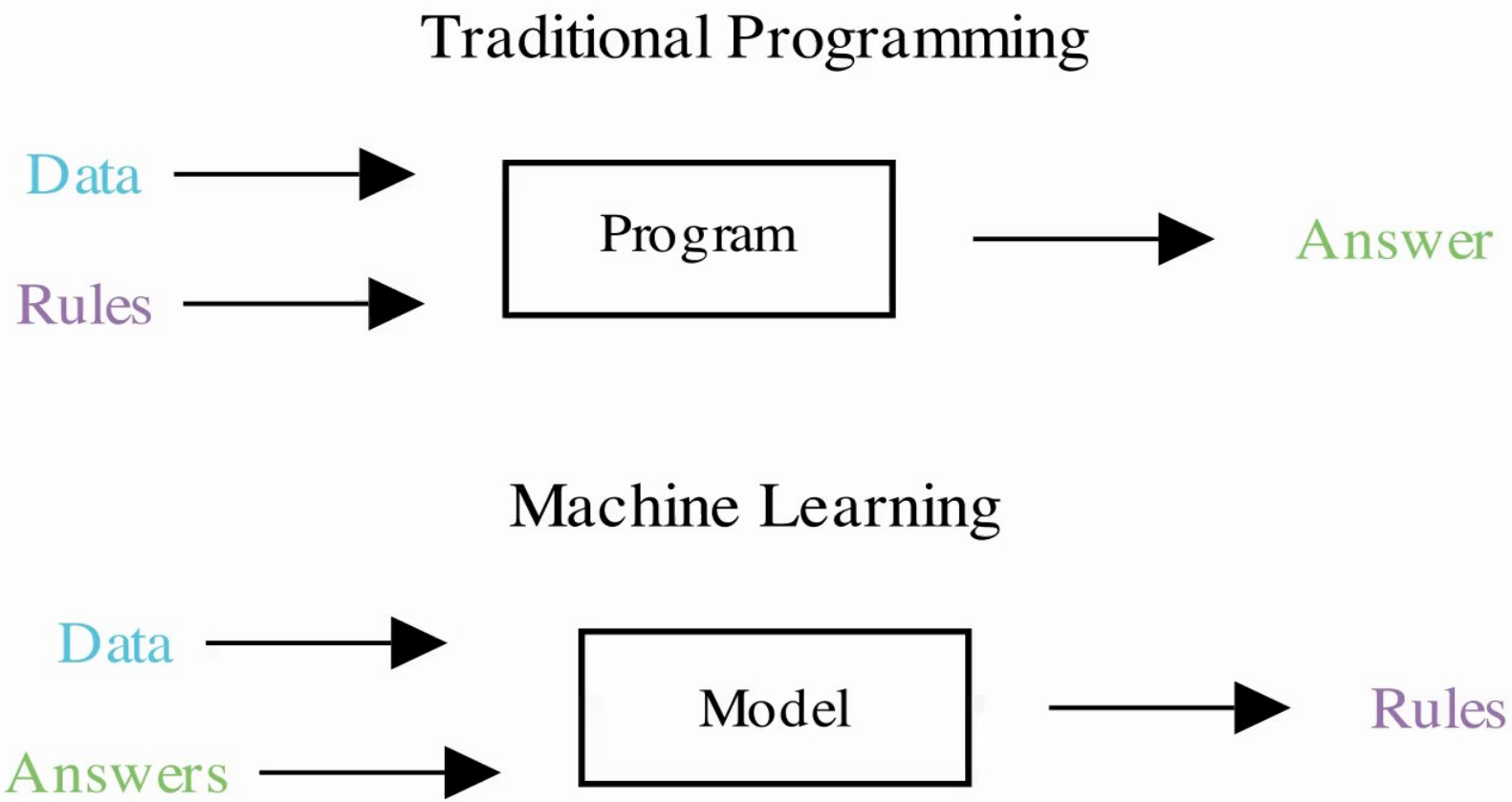


"Gradient"

AI, ML & DL



ML vs Traditional Programming



Main ML Paradigms

Supervised Learning: labeled data → e.g. classification, regression

Unsupervised Learning: unlabeled data → e.g. clustering, PCA

Reinforcement Learning: agent learns via environment interaction

Real Examples

- Supervised: spam detection, house price prediction
- Unsupervised: customer segmentation, anomaly detection
- Reinforcement: self-driving cars, AI in games

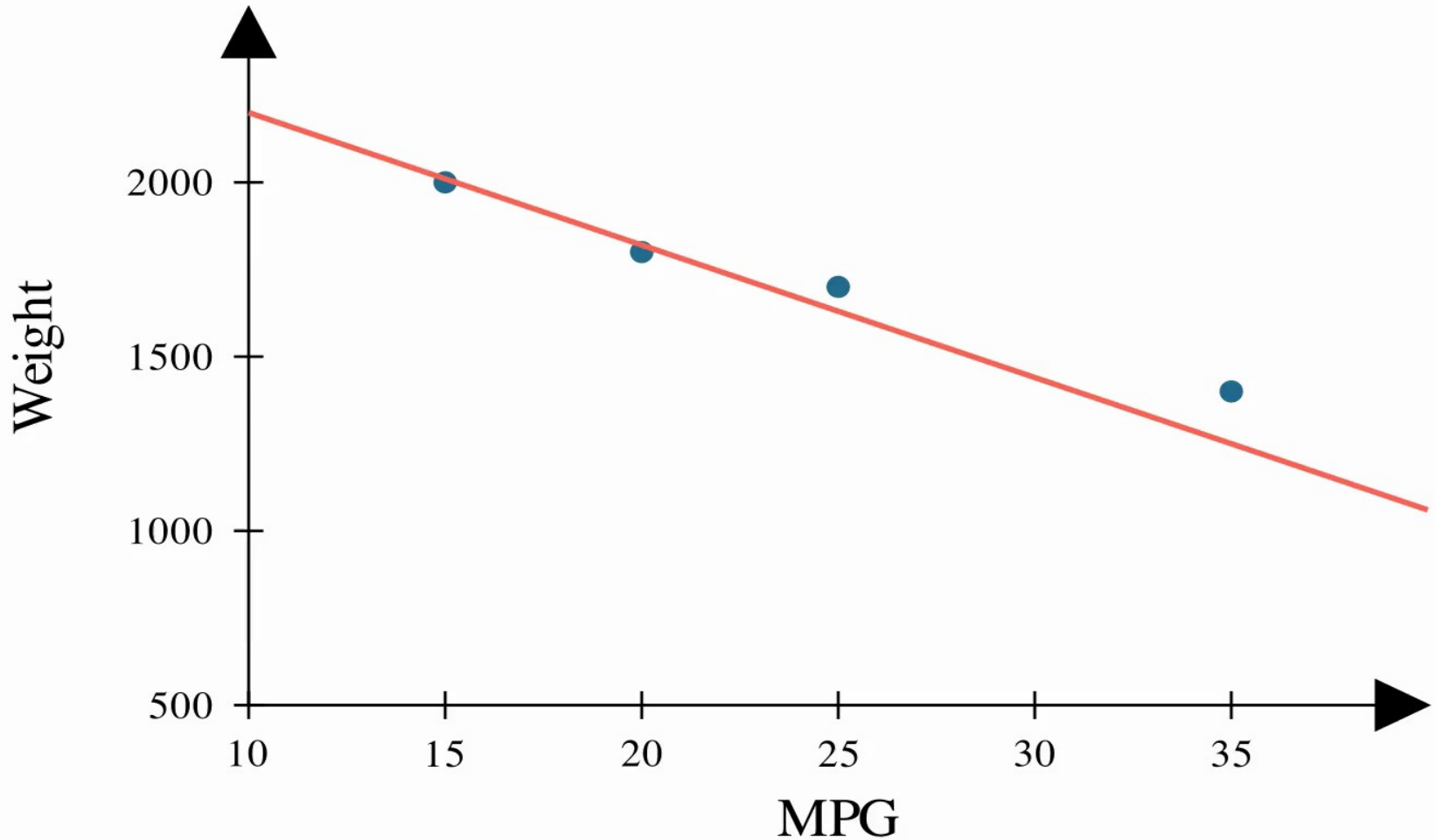
What is Regression?

Regression is a method to predict numerical values from data.

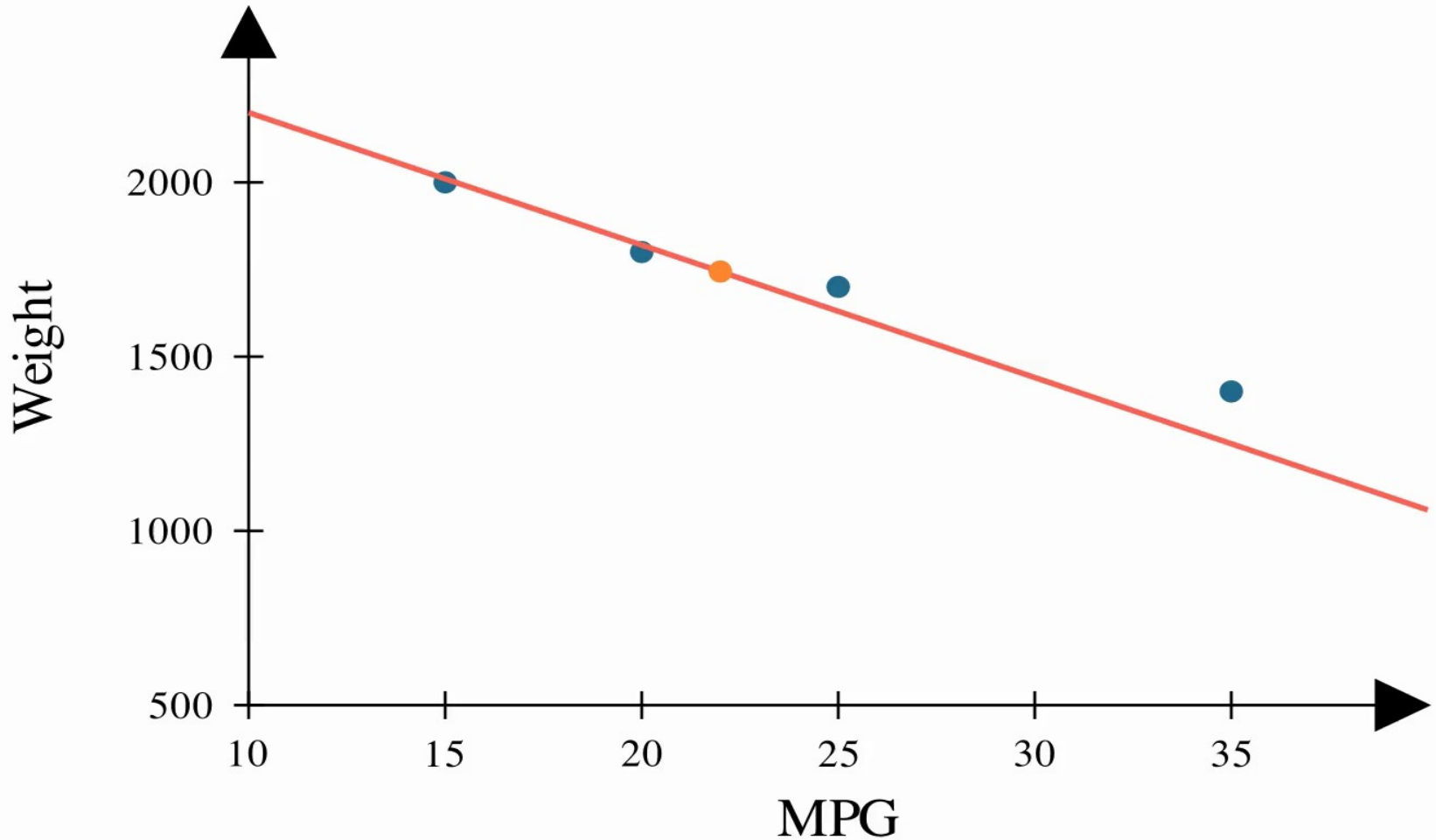
Examples of Regression

- Predicting car MPG
- Weather forecasting
- Trend analysis

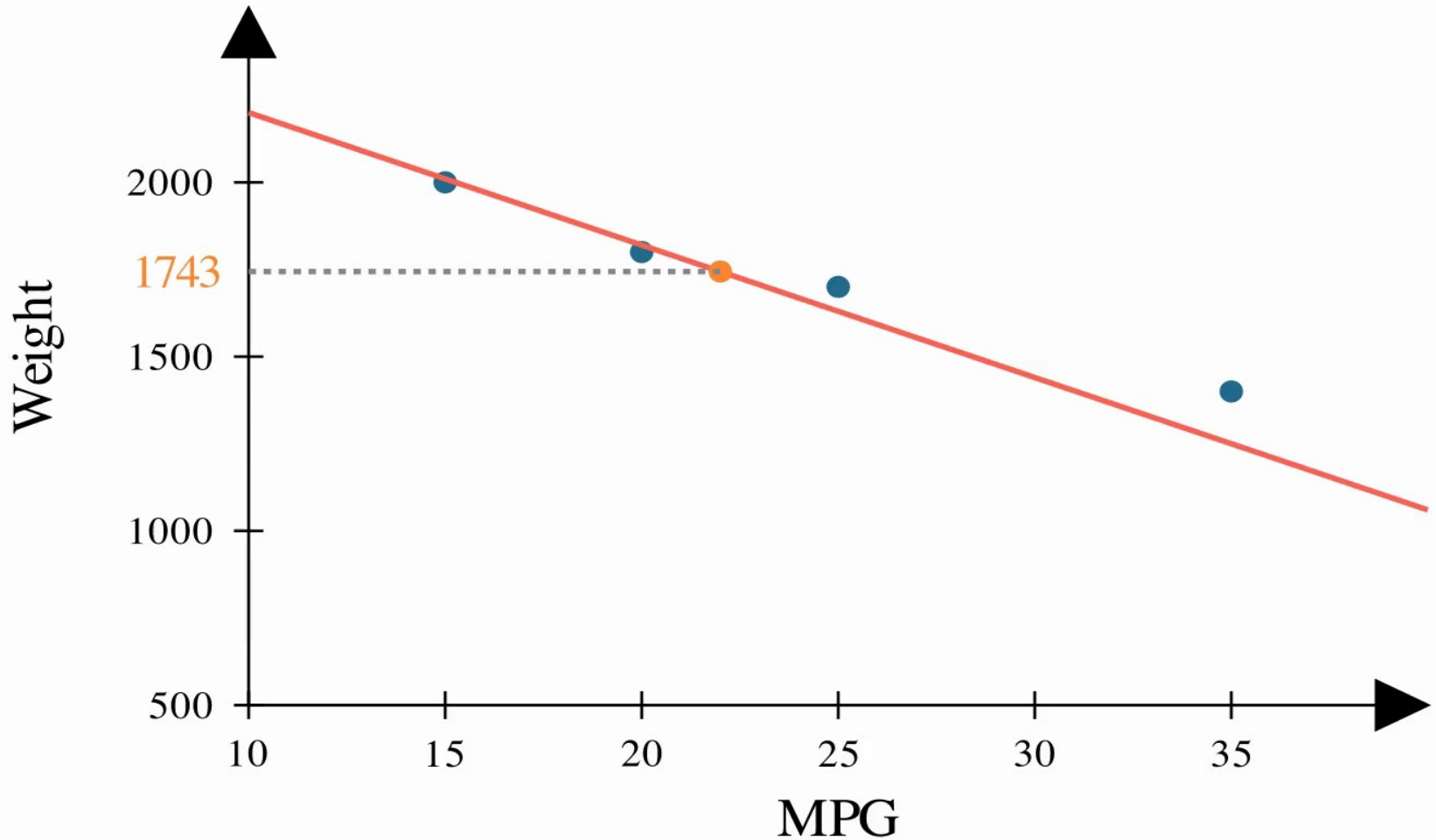
Linear Regression Example



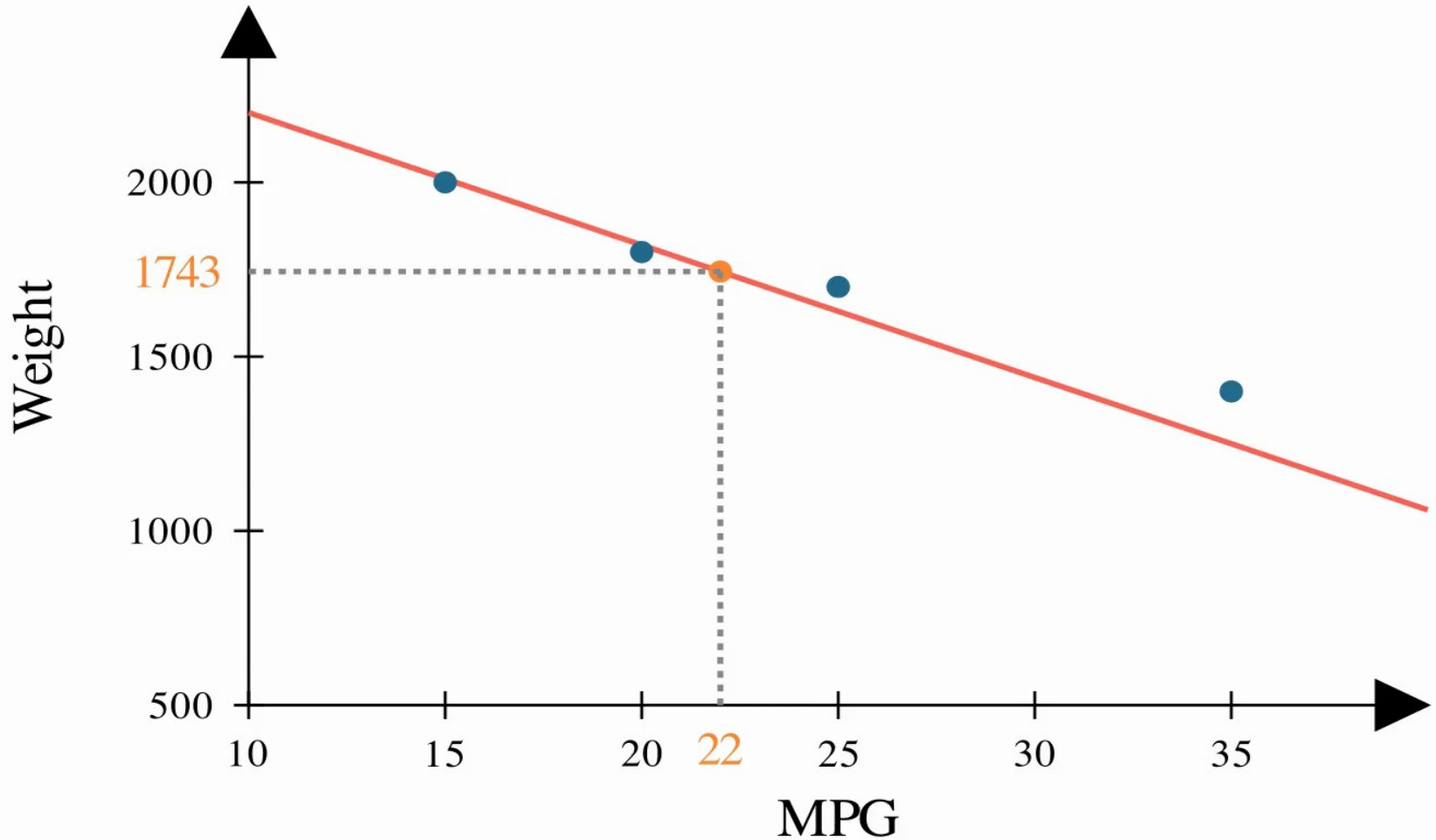
Linear Regression Example



Linear Regression Example



Linear Regression Example



Regression Formula

$$f(x) = a \cdot x + b$$

Regression Formula

$$\hat{y} = w \cdot x + b$$

Training Models

Training Models

Dataset

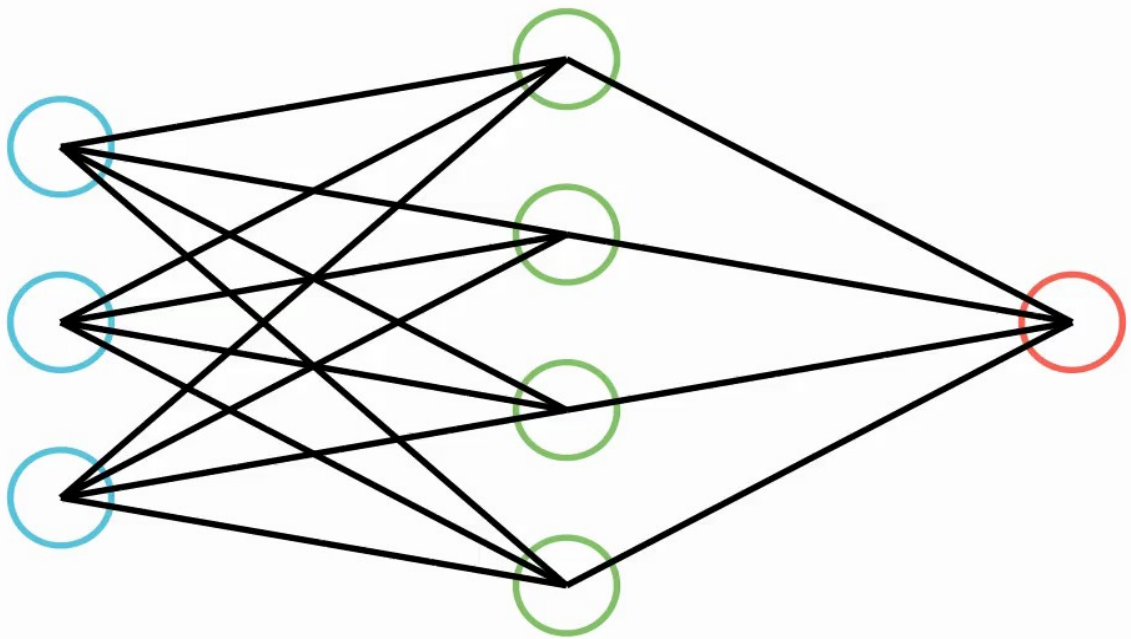
Train set	Validation set	Test set
≈70%	≈20%	≈10%

Training Models

Training Models

Model

$$\hat{y} = w \cdot x + b$$



Training Models

Loss

$$MSE = \frac{1}{n}\sum_{i}(y_i - \hat{y}_i)^2$$

Label: 2 Prediction: 2.1

Label: 3 Prediction: 2.9

Label: 4 Prediction: 4.2

$$MSE = \frac{1}{3}\left((2-2.1)^2 + (3-2.9)^2 + (4-4.2)^2 \right) = 0.0267$$

Training Models

Optimizer

Gradient Descent

$$\nabla_w L = \frac{1}{n}\sum_i 2((wx_i + b) - y_i)x_i, \quad \nabla_b L = \frac{1}{n}\sum_i 2((wx_i + b) - y_i)$$

$$g_{w,1} = 2 * (2.5 - 2) * 1 = 1.00$$

$$g_{b,1} = 2 * (2.5 - 2) = 1.00$$

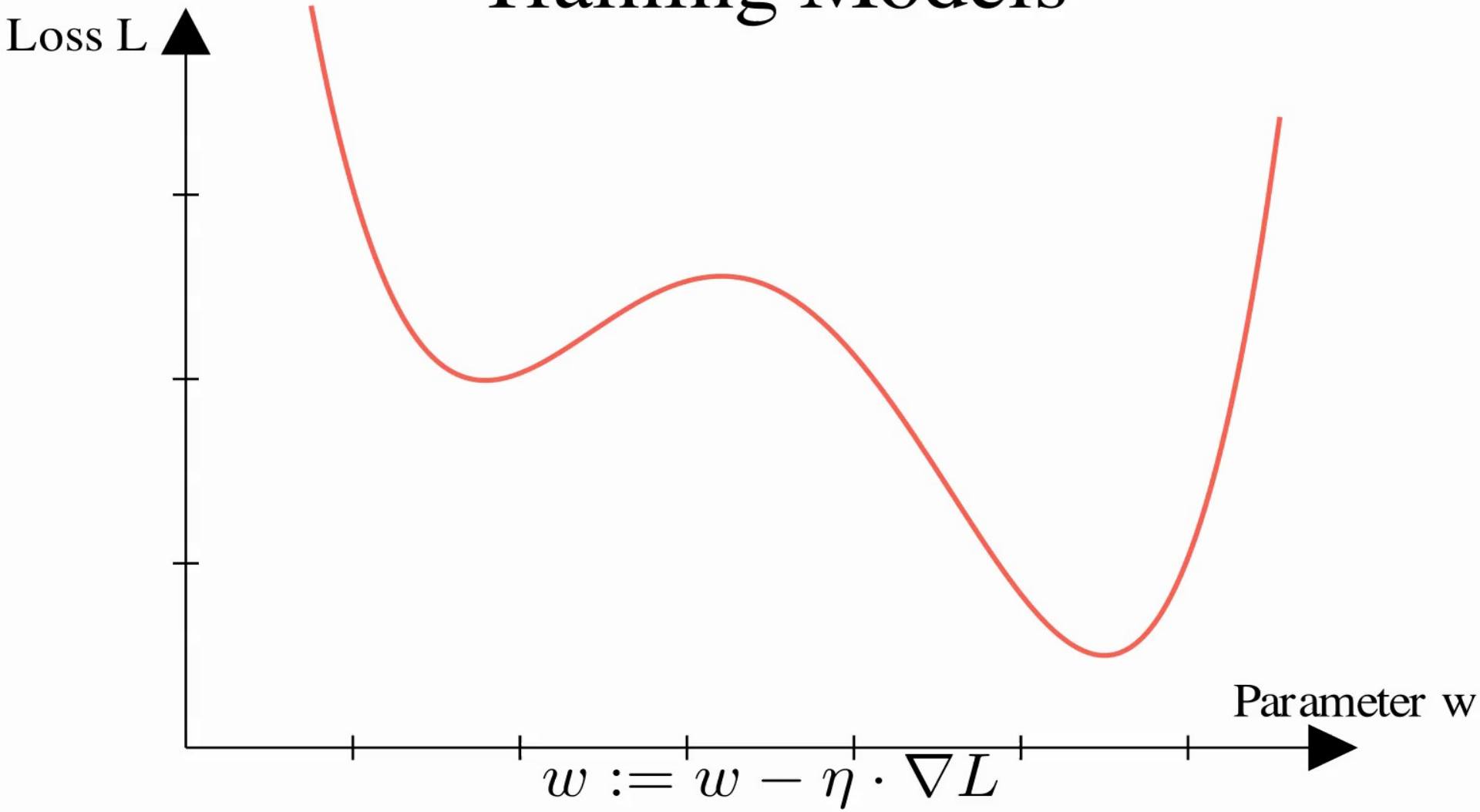
$$g_{w,2} = 2 * (2.8 - 3) * 2 = -0.80$$

$$g_{b,2} = 2 * (2.8 - 3) = -0.40$$

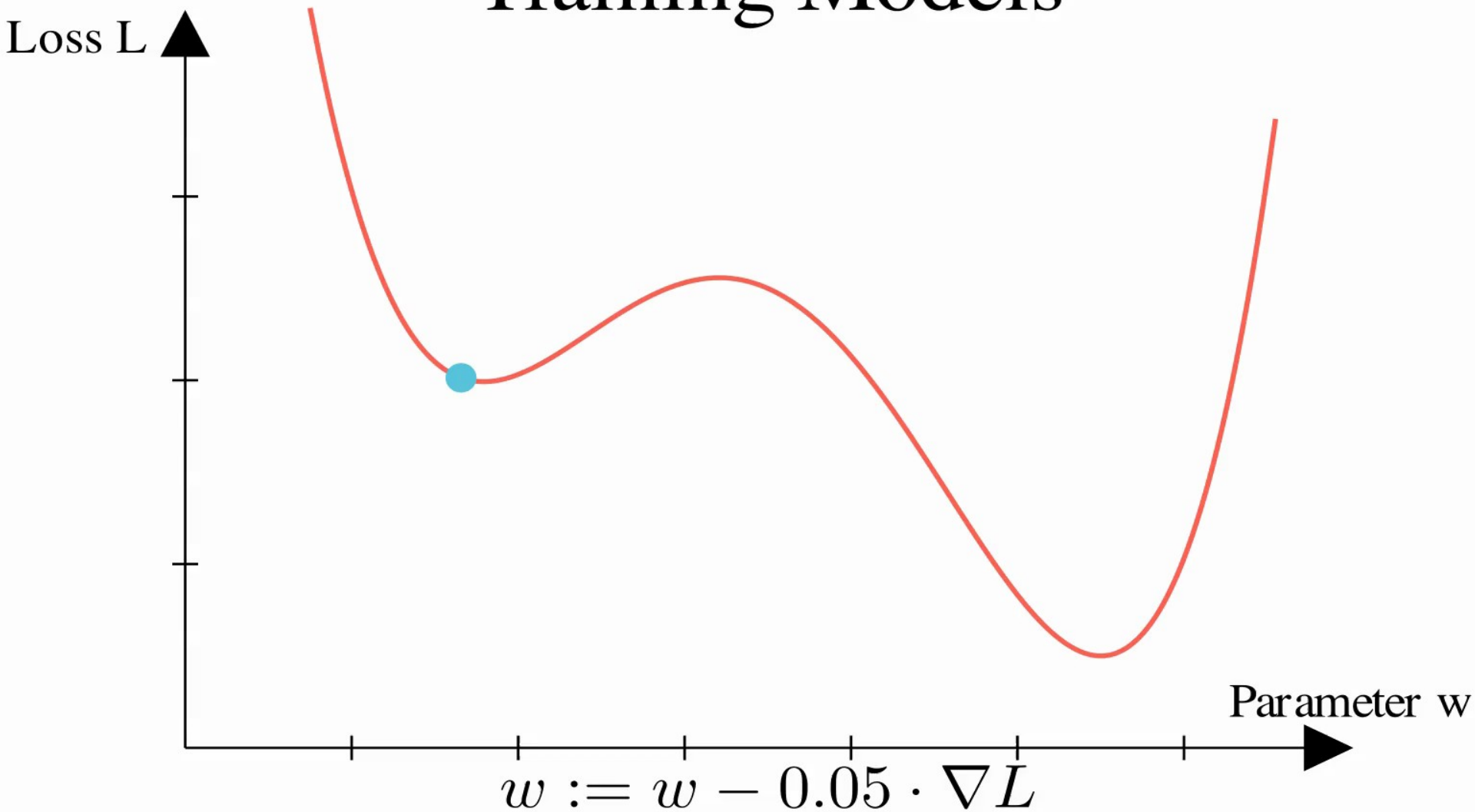
$$g_{w,avg} = \frac{1.00 + -0.80}{2} = 0.100$$

$$g_{b,avg} = \frac{1.00 + -0.40}{2} = 0.300$$

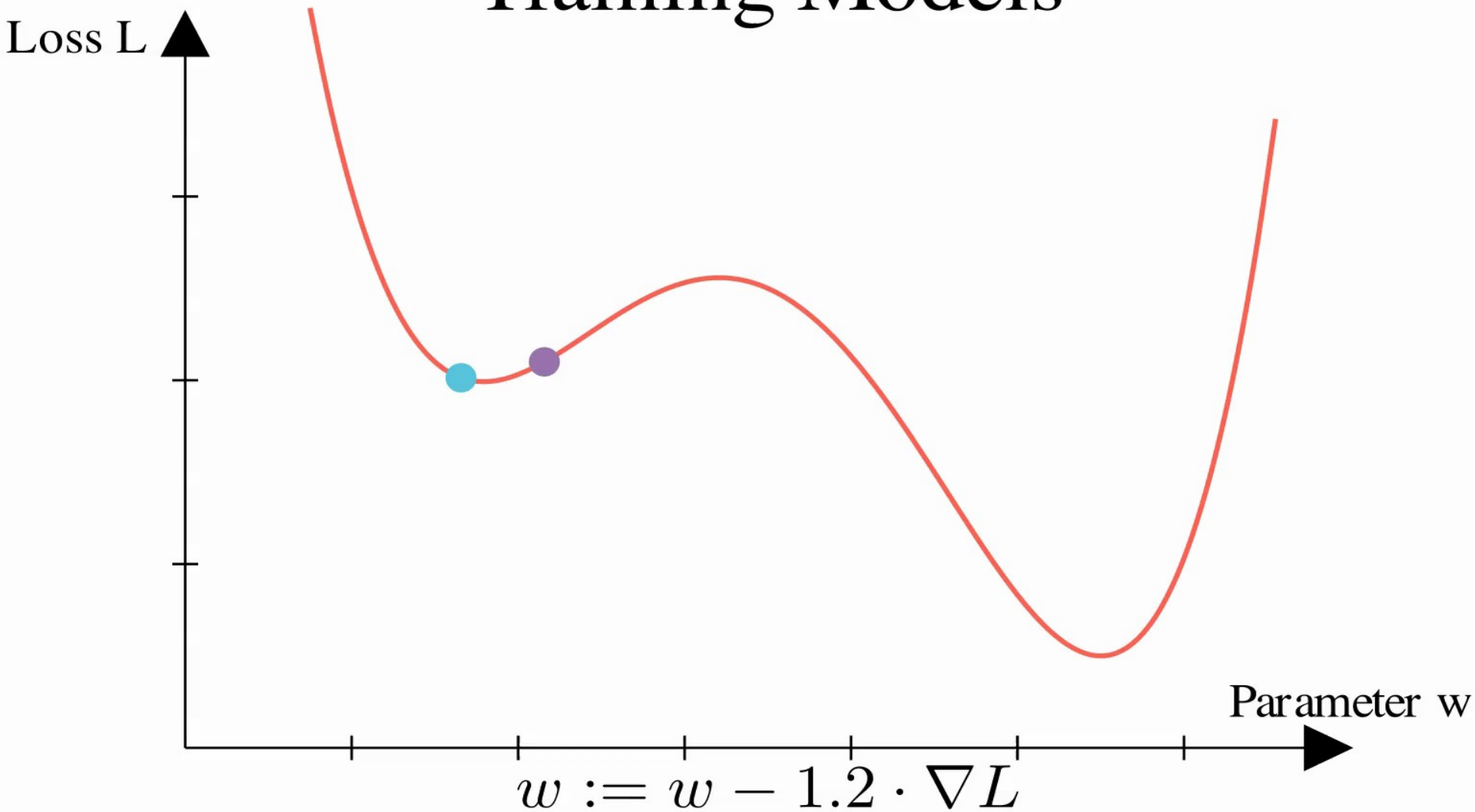
Training Models



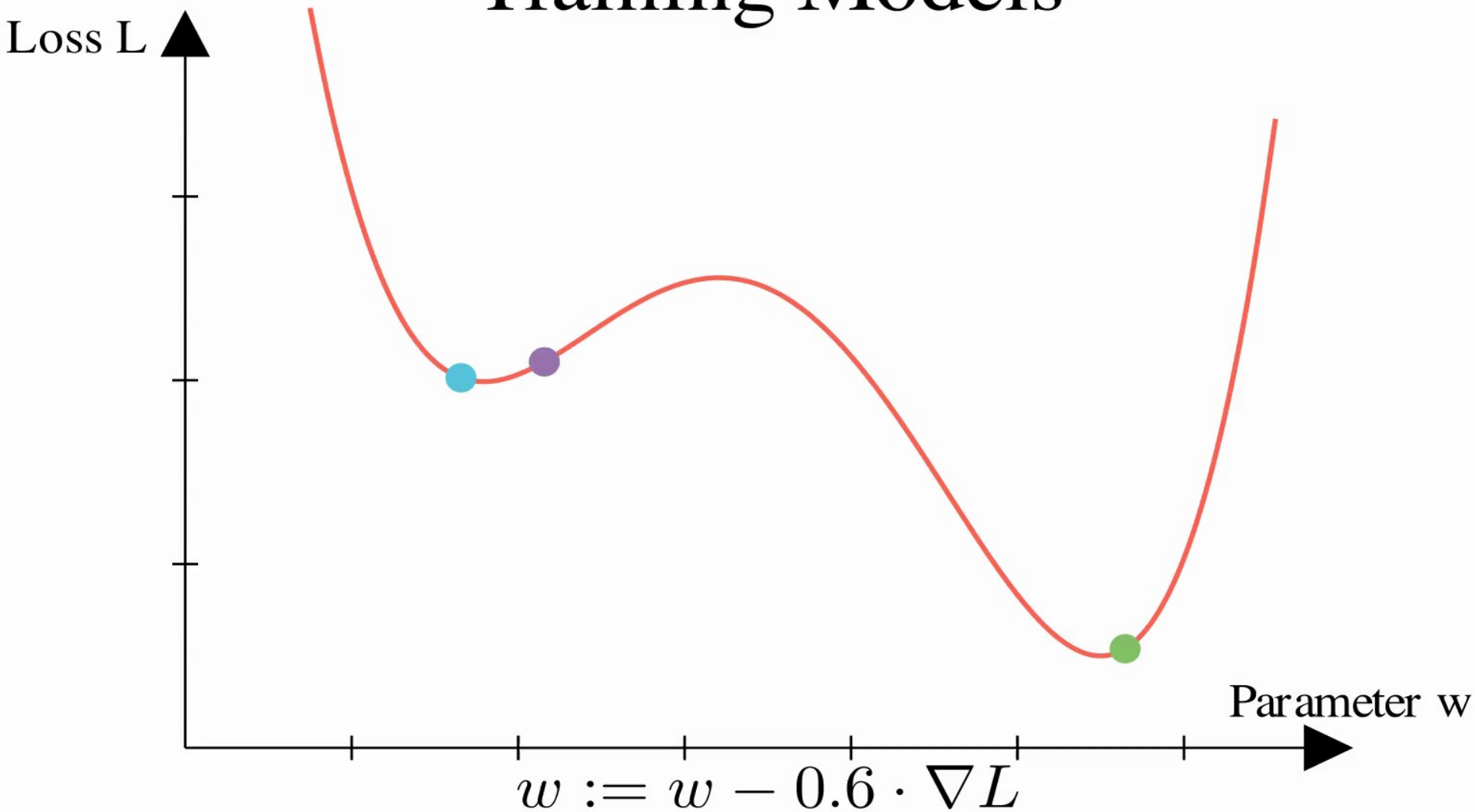
Training Models



Training Models

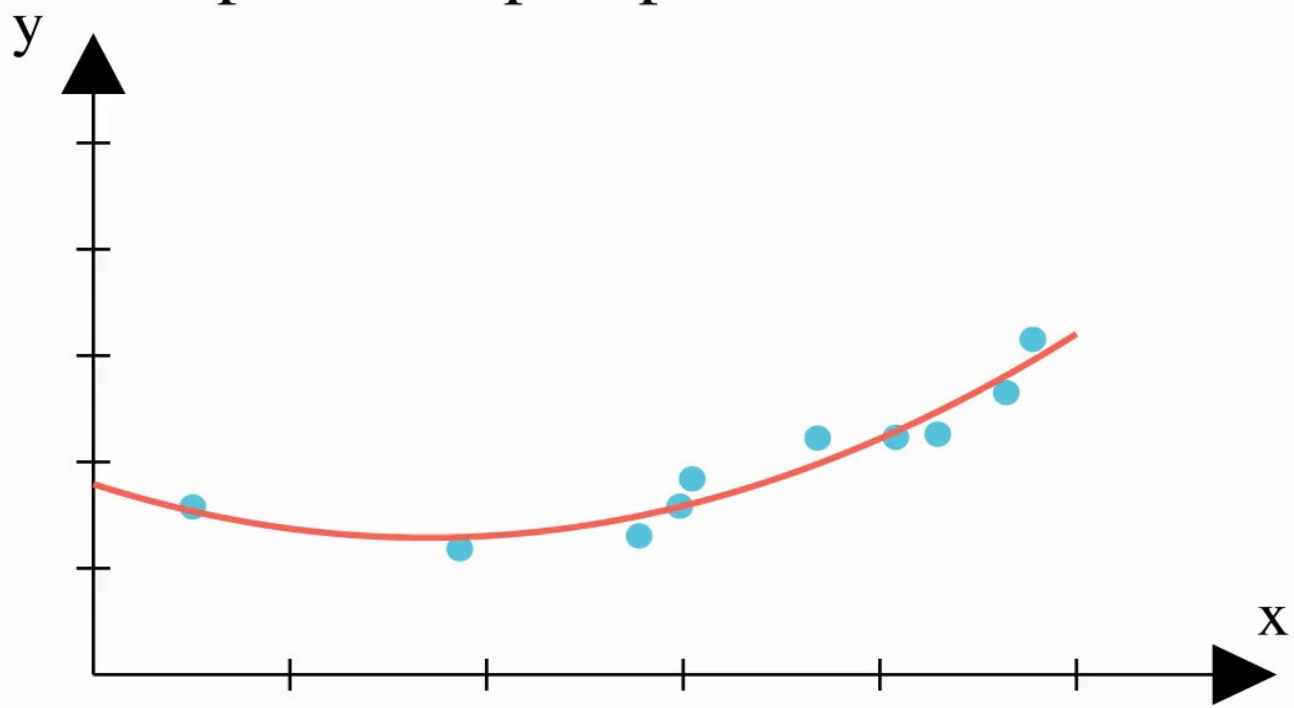


Training Models



Non-linear Regression

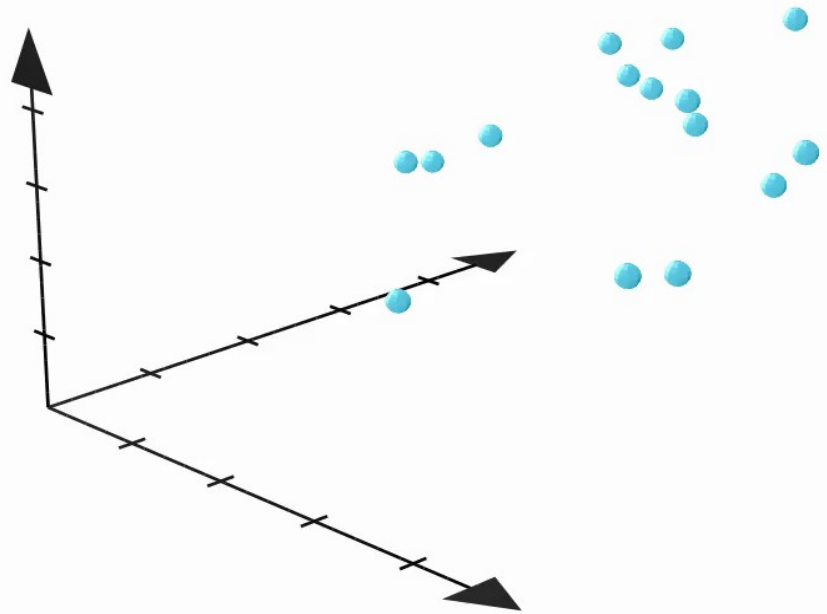
Used when the relationship between input and output is not linear.
Fits curves to capture complex patterns.



$$\hat{y} = \theta_0 + \theta_1 x + \theta_2 x^2$$

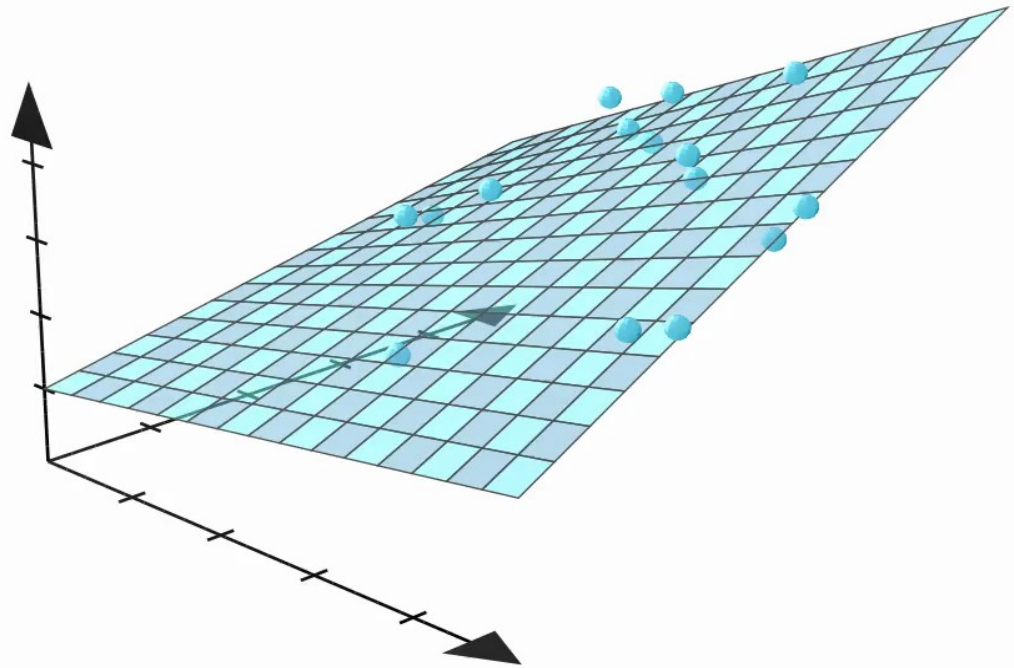
Multidimensional Regression

Used when the output depends on multiple input features.
Fits surfaces to capture complex patterns in higher dimensions.



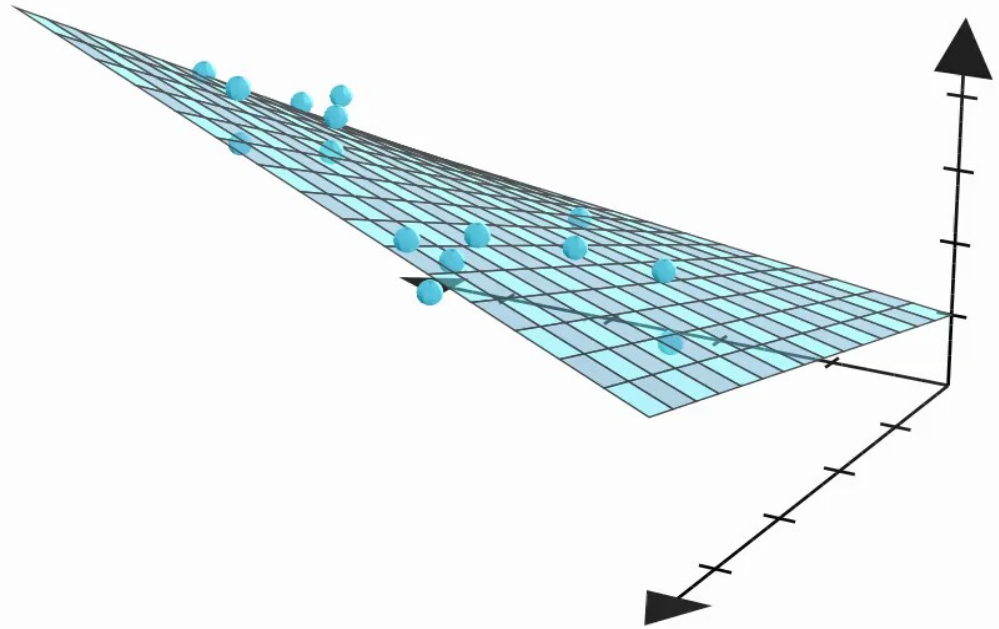
Multidimensional Regression

Used when the output depends on multiple input features.
Fits surfaces to capture complex patterns in higher dimensions.



Multidimensional Regression

Used when the output depends on multiple input features.
Fits surfaces to capture complex patterns in higher dimensions.



$$\hat{y} = \theta_0 + \theta_1 x + \theta_2 z + \theta_3 xz$$

Questions & Discussion

Thank you!