Machine Learning:

Machine learning is a field of artificial intelligence that teaches computers to learn from data without being explicitly programmed.

Types of Machine learning :

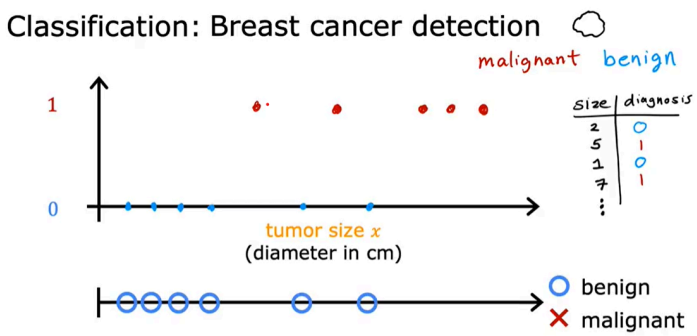
1. Supervised Learning
2. Unsupervised Learning
3. Semi-Supervised Learning
4. Reinforcement Learning

Supervised Learning :

In supervised learning, the algorithm learns from labeled data, meaning each training example has an input and a known correct output (label). The model makes predictions and adjusts based on the accuracy of those predictions.

Types of Supervised Learning Tasks:

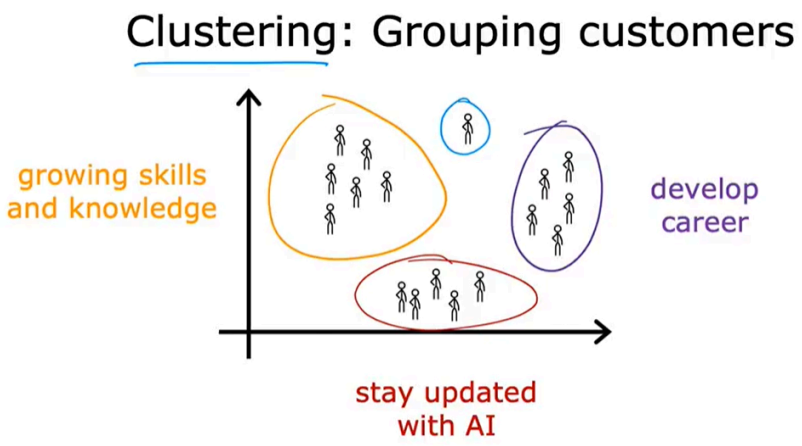
Task Type	Output Type	Example
Regression	Continuous Value	Predicting house prices, stock value
Classification	Discrete Class Label	Spam detection



Unsupervised Learning :

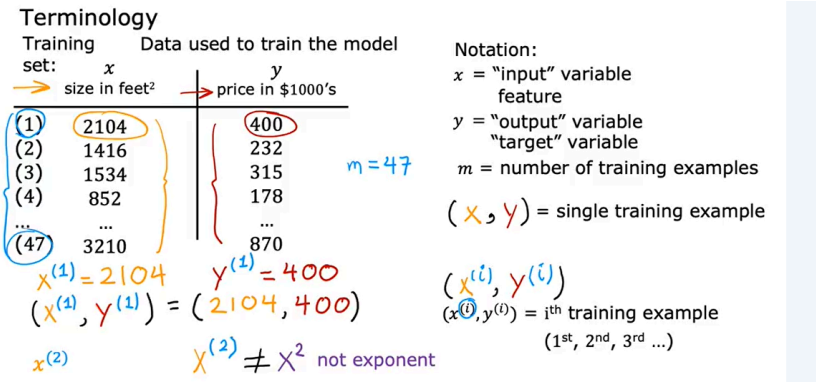
Unsupervised learning deals with unlabeled data; the model explores data structure or patterns without predefined labels.

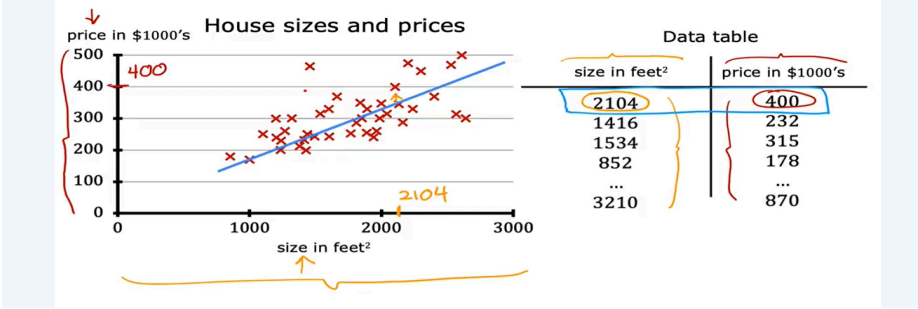
Main applications: Clustering, dimensionality reduction, anomaly detection.



Linear Regression Model:

Linear Regression is a supervised learning algorithm used to predict a continuous output variable based on one or more input features. It models the relationship between the input variables (X) and the output variable (y) using a straight line.

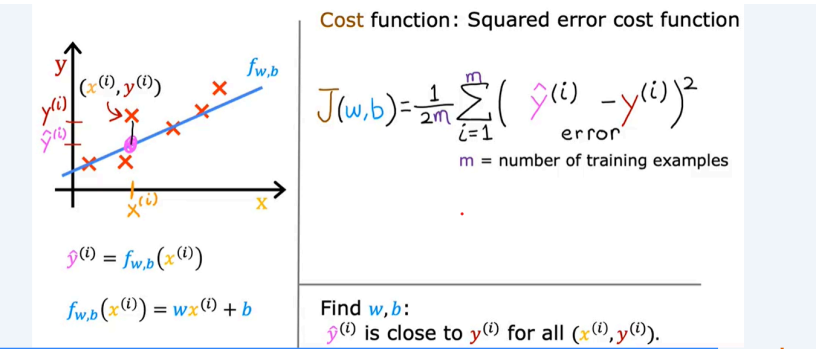


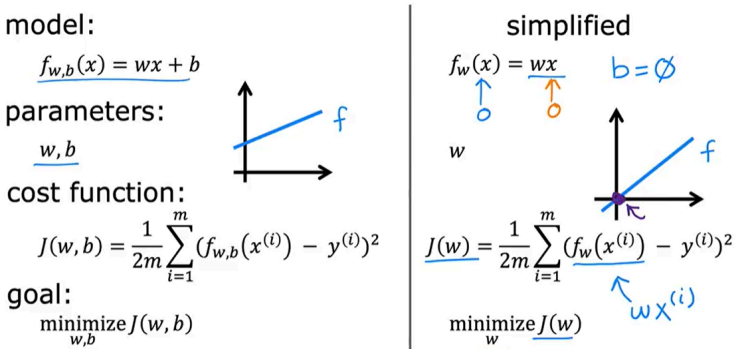


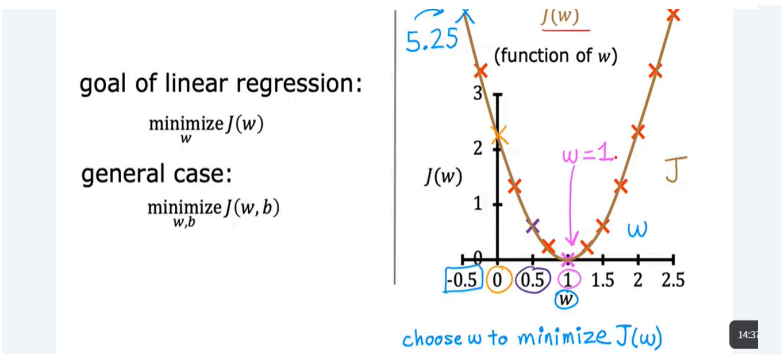
Cost Function:

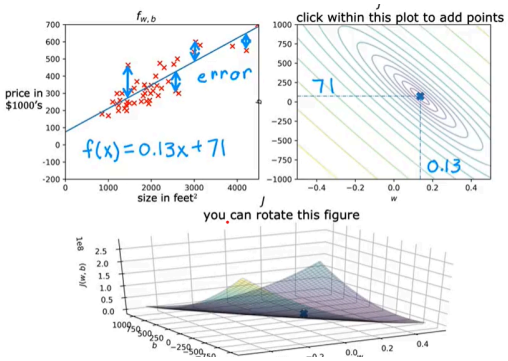
It calculates the difference between predicted values and actual values from the dataset.

🎯 **Goal:** Minimize the cost function to improve model accuracy.



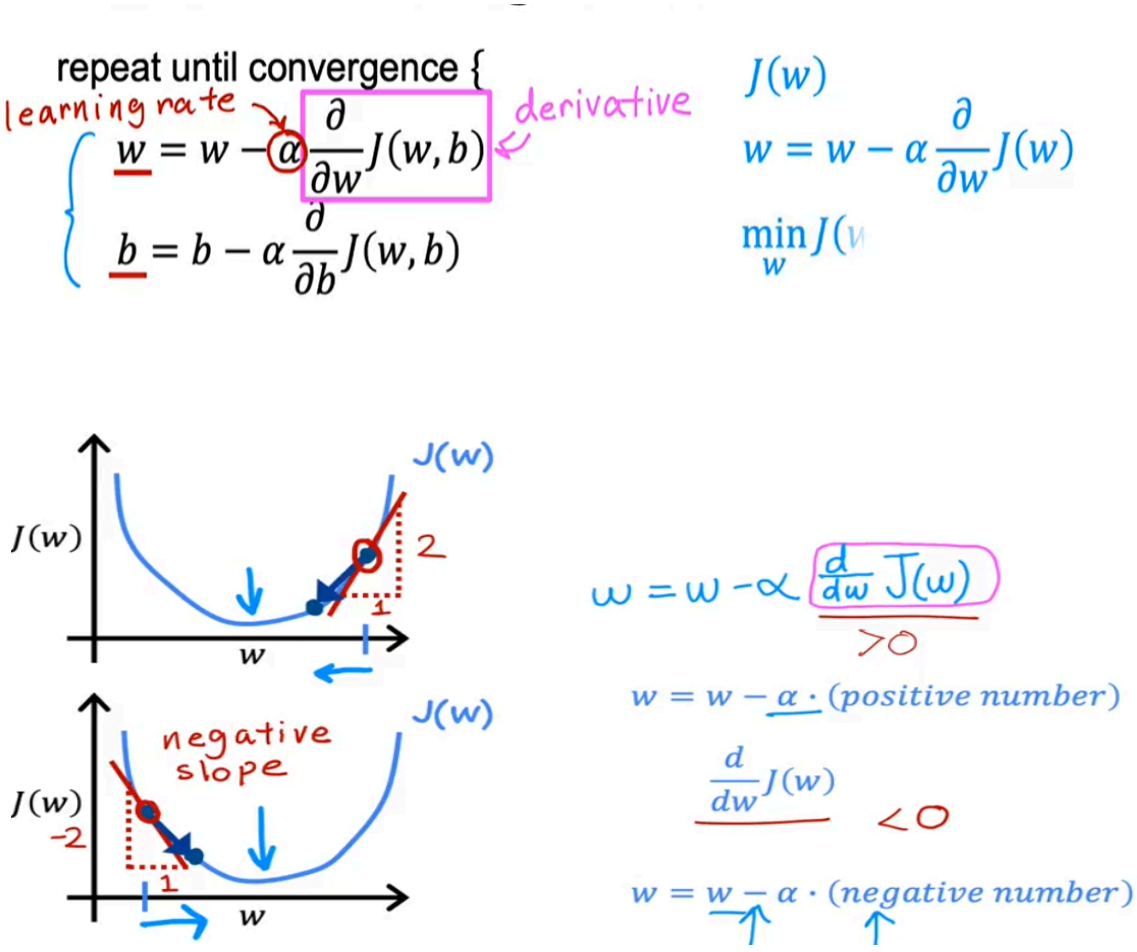




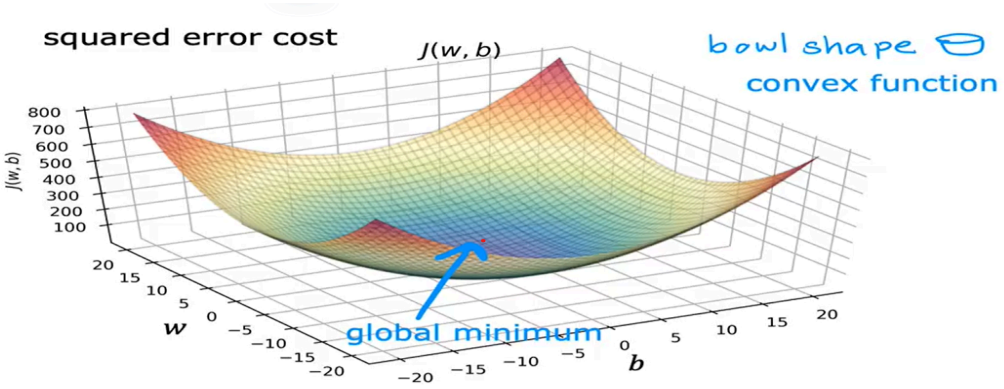


Gradient Descent:

Gradient Descent is an optimization algorithm used to minimize the cost function (or loss function) in machine learning models, typically through iterative updates to model parameters.



Gradient descent for linear regression



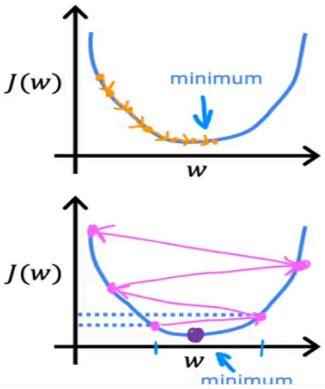
Learning Rate:

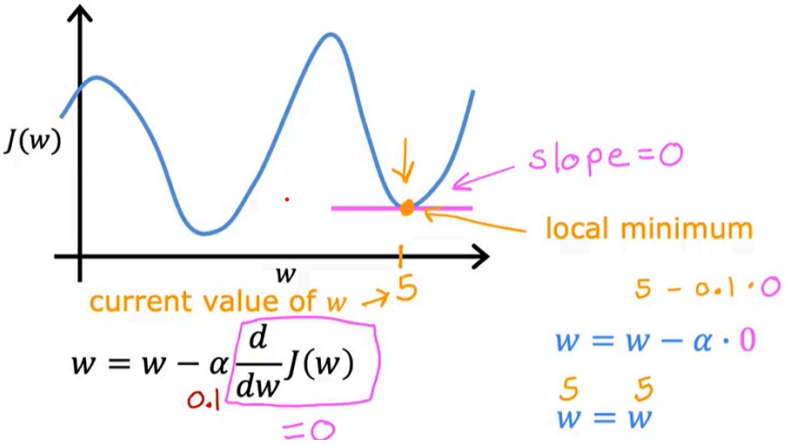
The learning rate, often denoted by α (alpha), is a hyperparameter that determines the size of the steps taken during gradient descent when updating model parameters.

$$w = w - \alpha \frac{d}{dw} J(w)$$

If α is too small...
Gradient descent may be slow.

If α is too large...
Gradient descent may:
- Overshoot, never reach minimum





Can reach local minimum with fixed learning rate α

$$w = w - \alpha \frac{d}{dw} J(w)$$

smaller
not as large
large

Near a local minimum,
- Derivative becomes smaller
- Update steps become smaller

Can reach minimum without decreasing learning rate α