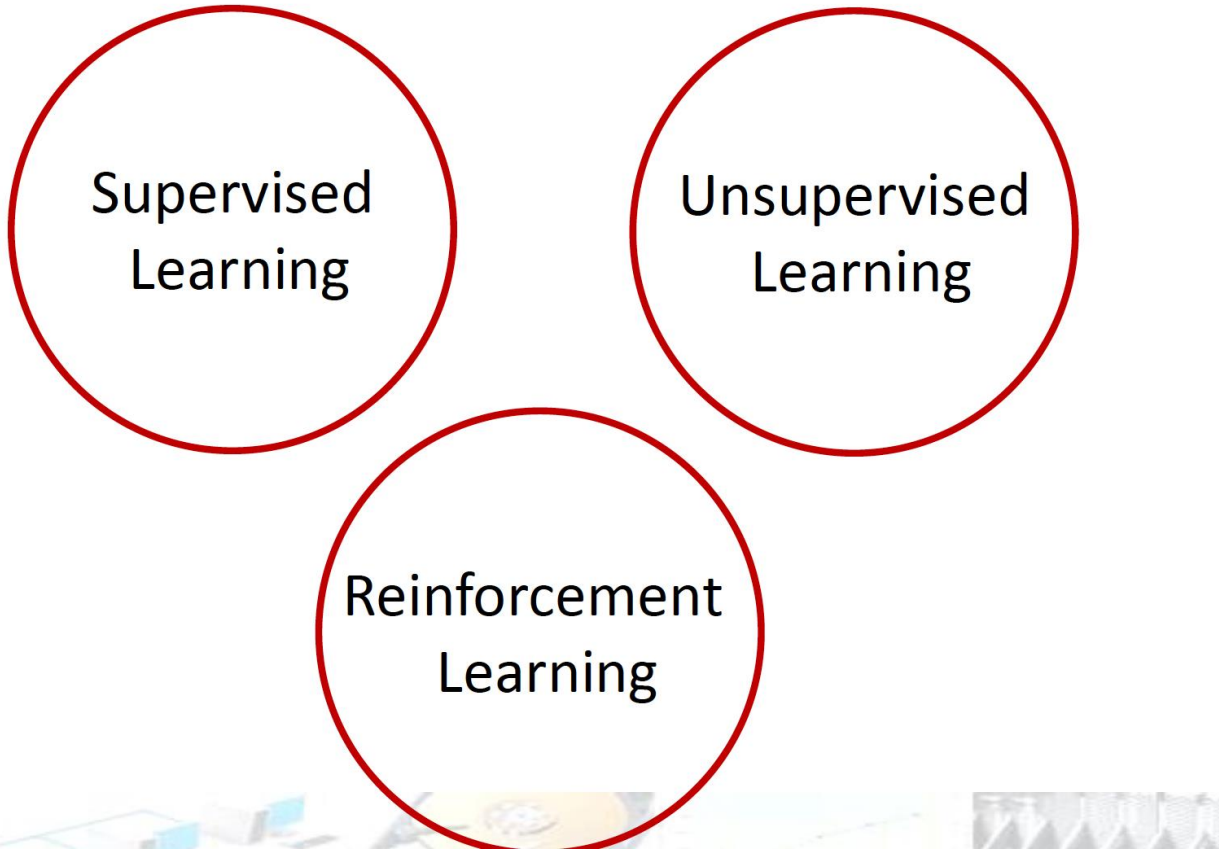
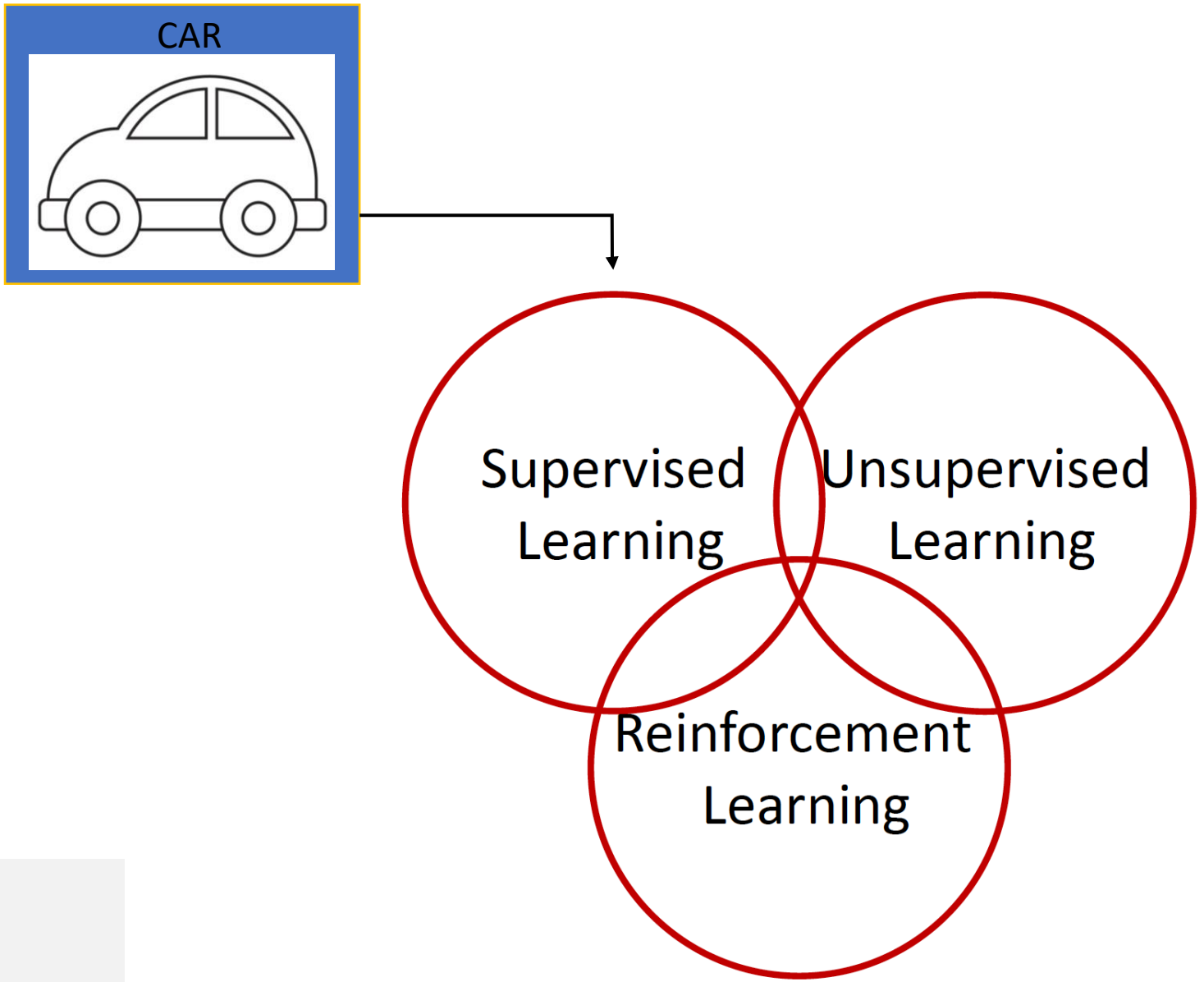
Types of Machine Learning

Types of Machine Learning



Types of Machine Learning



Types of Learning

Supervised (inductive) learning

- Given: training data + desired outputs (labels)

Unsupervised learning

- Given: training data (without desired outputs)

Semi-supervised learning

- Given: training data + a few desired outputs

Types of Machine Learning

Reinforcement learning

- A reinforcement learning algorithm, or agent, learns by interacting with its environment.
- The agent receives rewards by performing correctly and penalties for performing incorrectly.
- The agent learns without intervention from a human by maximizing its reward and minimizing its penalty.

Supervised Learning

We are given input samples (X) and output samples (y) of a function y = f(X).

We can represent the entire data set:

Data= (X,y); {Standard Notation}

We would like to “learn” f, and evaluate it on new data. Types:

–Classification: y is factors (class labels).

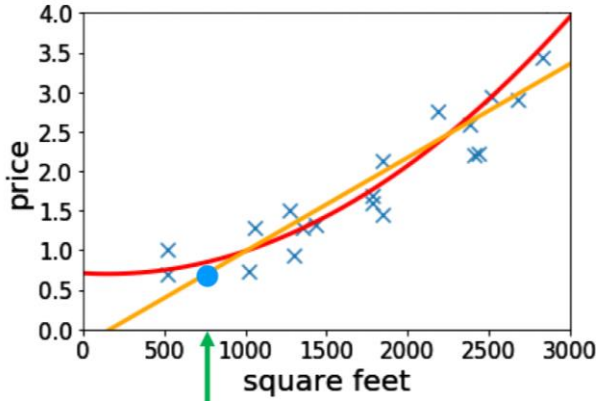
–Regression: y is continuous, e.g. linear regression.

Housing Price Prediction

- Given: a dataset that contains n samples

$$(x^{(1)}, y^{(1)}), \ldots (x^{(n)}, y^{(n)})$$

- Task: if a residence has x square feet, predict its price?

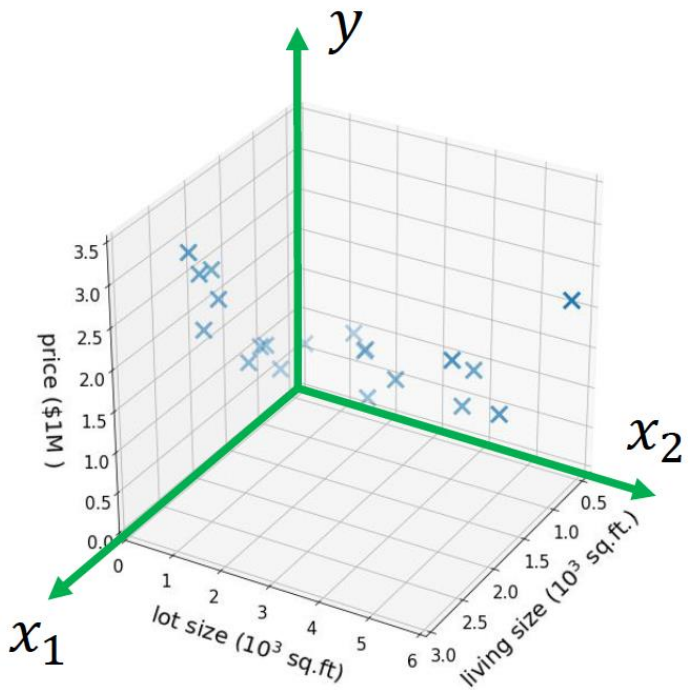


- fitting linear/qaudratic functions to the dataset

X= 800

Y= ?

More Features



- Suppose we also know the lot size
- Task: find a function that maps

(size, lot size) → price

features/input $x \in \mathbb{R}^2$ label/output $y \in \mathbb{R}$

- Dataset: $\left(x^{(1)}, y^{(1)}\right), \ldots, (x^{(n)}, y^{(n)})$

where $x^{(i)} = (x_1^{(i)}, x_2^{(i)})$

- "Supervision" refers to $y^{(1)}, \ldots, y^{(n)}$

High Dimensional Features

- $x \in \mathbb{R}^d$ for large d
- E.g.,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ \vdots \\ \vdots \\ x_d \end{bmatrix}$$

--- living size
--- lot size
--- # floors
--- condition
--- zip code
⋮

→ y= price

Unsupervised Leaning

Given only samples X of the data (unlabelled data), we compute a function used to draw inferences.

–y is factor: Clustering

–y is continuous: Matrix factorization, Kalman filtering, unsupervised neural networks.