Supervised vs. Unsupervised

Given: training data

example $\vec{x_1} \rightarrow$	x_{11}	x_{12}	$\ldots$	x_{1d}	$y_1 \leftarrow$ label
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
example $\vec{x_i} \rightarrow$	x_{i1}	x_{i2}	$\ldots$	x_{id}	$y_i \leftarrow$ label
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
example $\vec{x_n} \rightarrow$	x_{n1}	x_{n2}	$\ldots$	x_{nd}	$y_n \leftarrow$ label

Supervised vs. Unsupervised

Given: training data

example $\vec{x_1} \rightarrow$	x_{11}	x_{12}	$\ldots$	x_{1d}	$y_1 \leftarrow$ label
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
example $\vec{x_i} \rightarrow$	x_{i1}	x_{i2}	$\ldots$	x_{id}	$y_i \leftarrow$ label
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
example $\vec{x_n} \rightarrow$	x_{n1}	x_{n2}	$\ldots$	x_{nd}	$y_n \leftarrow$ label

$$(\vec{x_1}, y_1), \ldots, (\vec{x_n}, y_n) \ / \ \vec{x_i} \in \mathbb{R}^d \text{ and } y_i \text{ is the label.}$$

Supervised vs. Unsupervised

fruit	length	width	weight	label
fruit 1	165	38	172	Banana
fruit 2	218	39	230	Banana
fruit 3	76	80	145	Orange
fruit 4	145	35	150	Banana
fruit 5	90	88	160	Orange
...				
fruit n	...	...	...	...

Supervised vs. Unsupervised

fruit	length	width	weight	label
fruit 1	165	38	172	Banana
fruit 2	218	39	230	Banana
fruit 3	76	80	145	Orange
fruit 4	145	35	150	Banana
fruit 5	90	88	160	Orange
...				
fruit n	...	...	...	...

Unsupervised learning

Learning a model from **unlabeled** data

Supervised learning

Learning a model from **labeled** data

Unsupervised Learning

Training data: "examples" $\vec{x}$

$$\vec{x_1}, \ldots, \vec{x_n}, \ \vec{x_i} \in X \subset \mathbb{R}^n$$

- **Clustering/segmentation:**

  $$f : \mathbb{R}^d \longrightarrow \{C_1, \ldots C_k\} \ \text{(set of clusters)}$$

 Example: Find clusters in the population, fruits, and species.

Unsupervised Learning
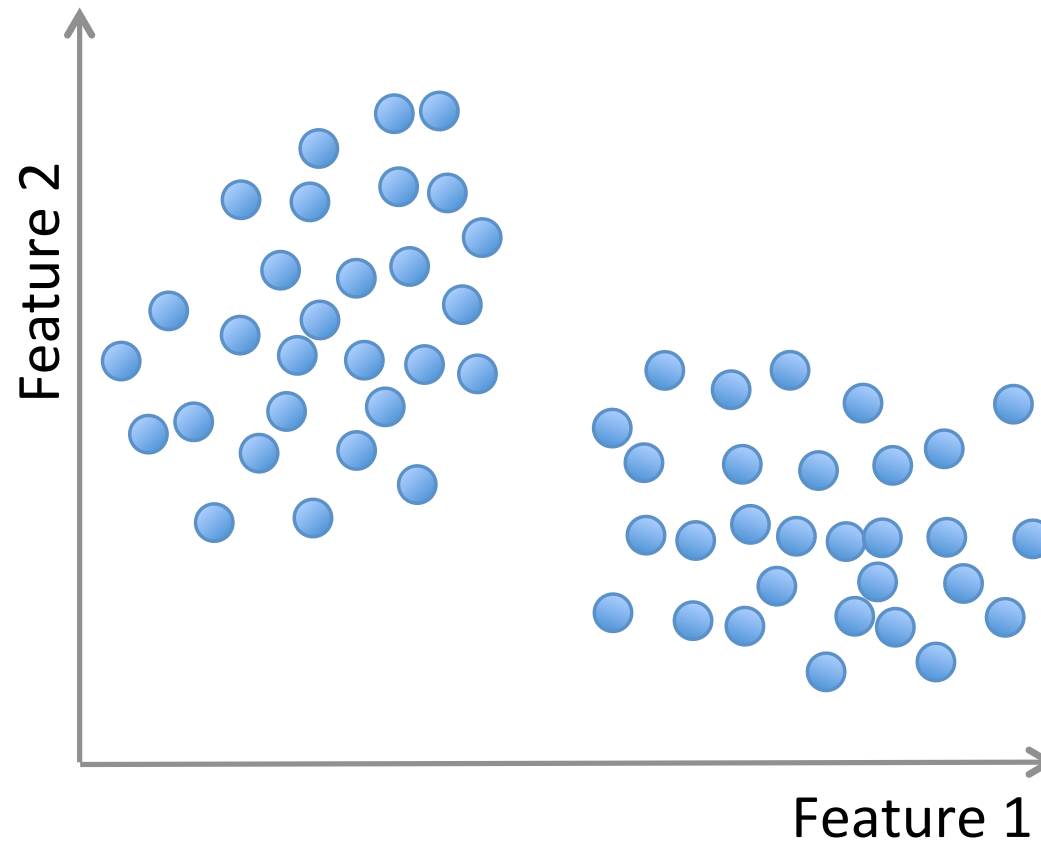
Unsupervised Learning
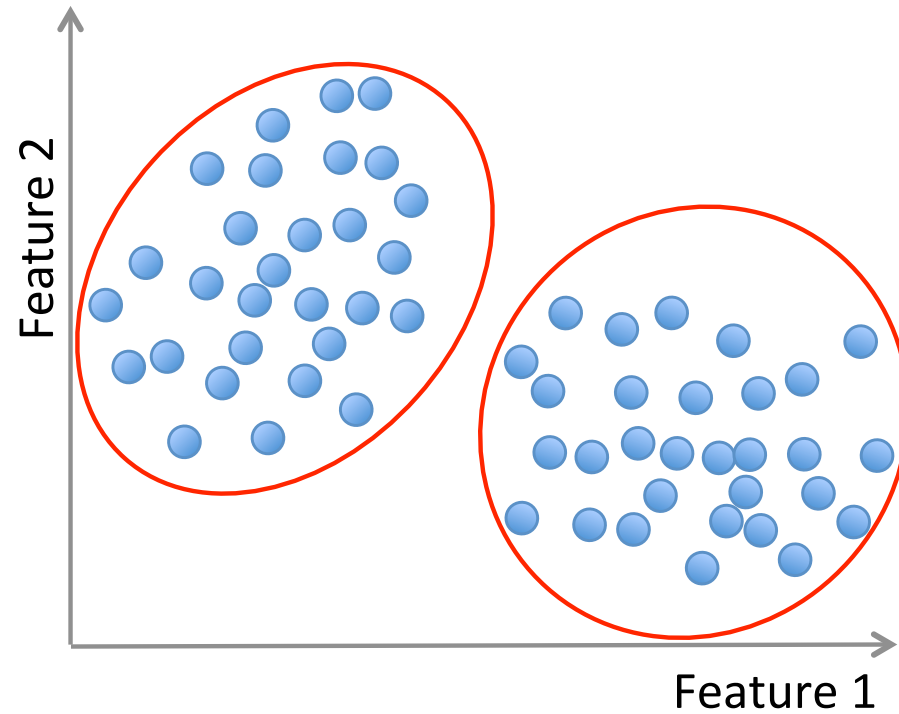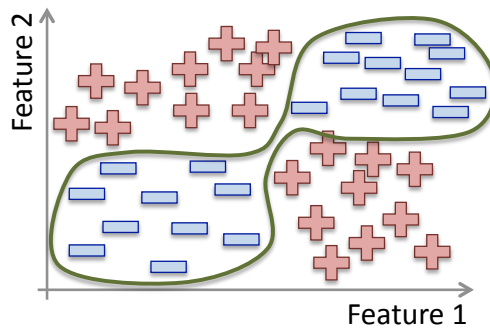
Unsupervised Learning
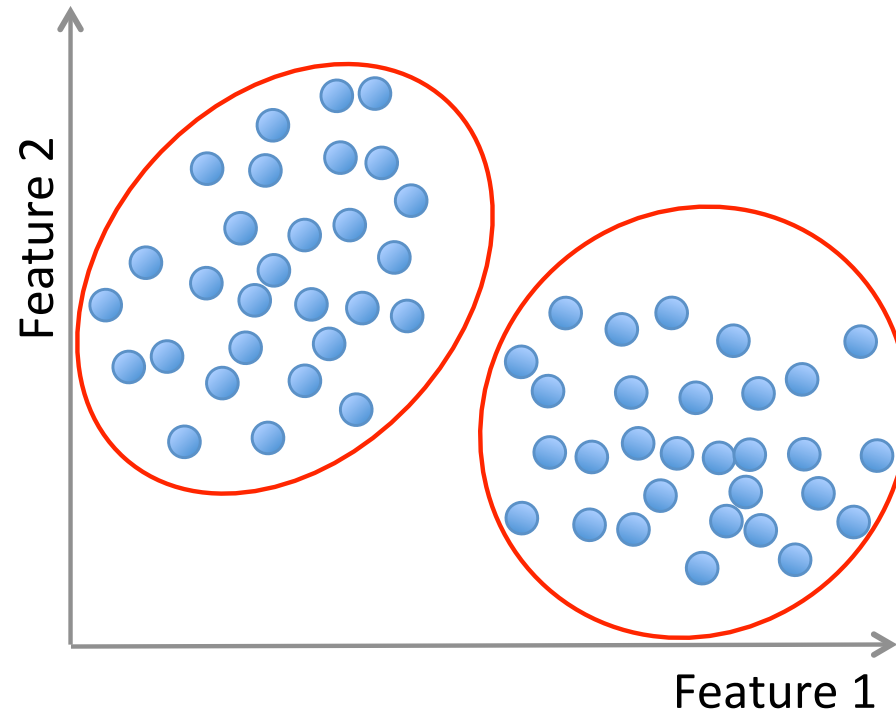
Methods: K-means, Gaussian mixtures, hierarchical clustering, spectral clustering, etc.

Supervised Learning

Training data: "examples" $\vec{x}$ with "labels" y

$$(\vec{x_1}, y_1), \ldots, (\vec{x_n}, y_n) \ / \ \vec{x_i} \in \mathbb{R}^d$$

- **Classification:** y is discrete; to simplify, $y \in \{-1, +1\}$

$$f : \mathbb{R}^d \longrightarrow \{-1, +1\} \qquad f \text{ is called a } \textbf{binary classifier}.$$

Example: approve credit yes/no, spam/ham, banana/orange
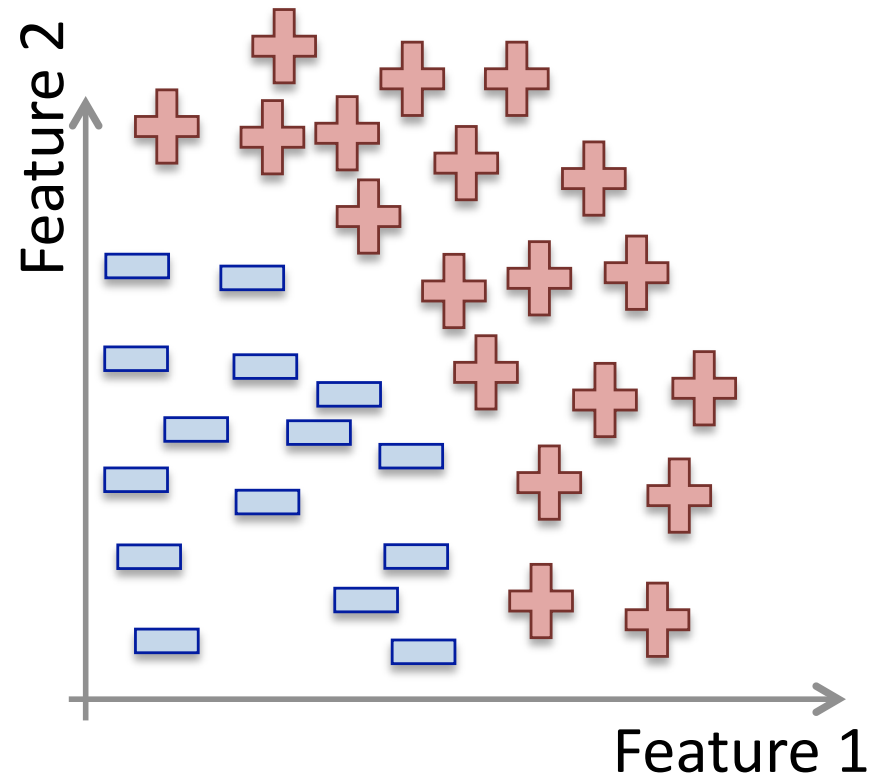
Supervised Learning

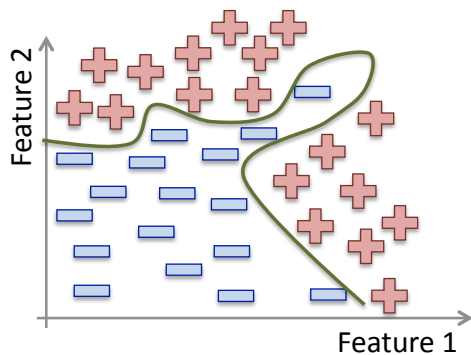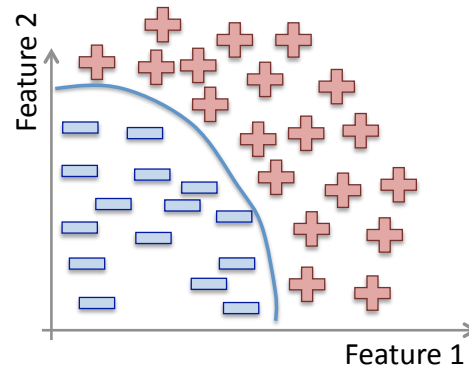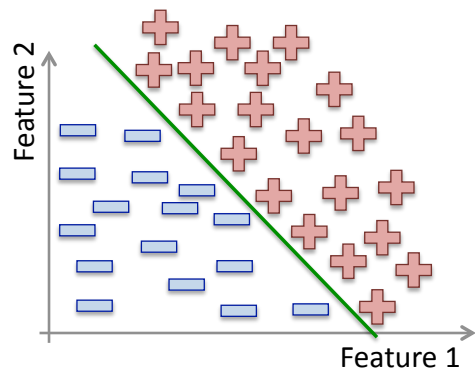# Supervised Learning

©**aiphabet**, Inc.
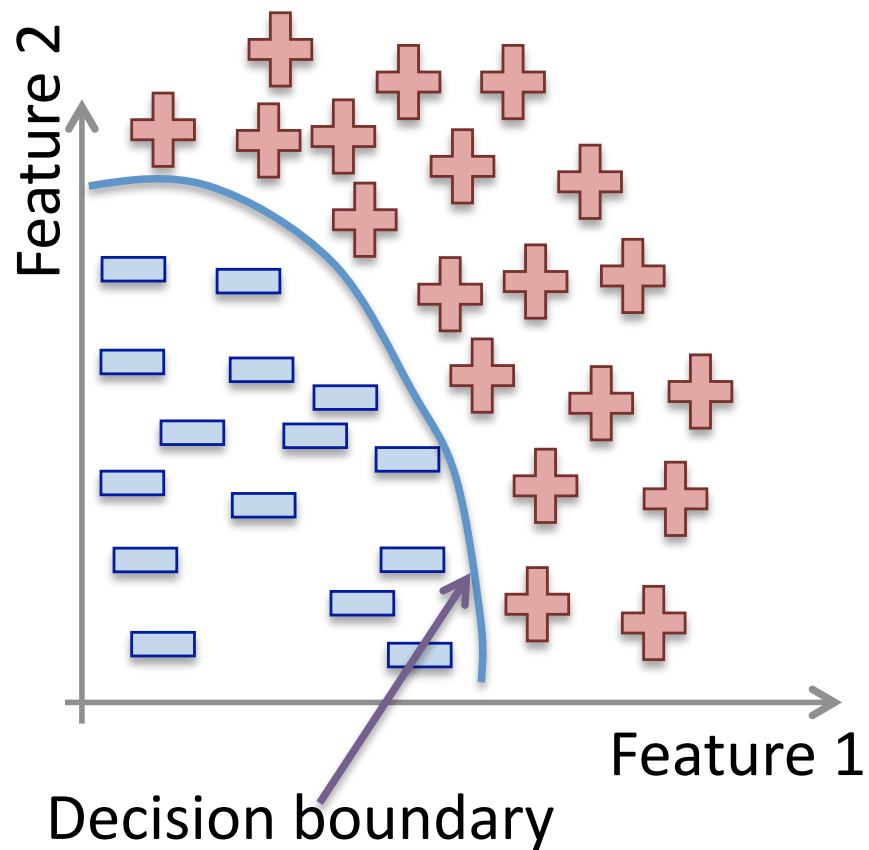
Supervised Learning
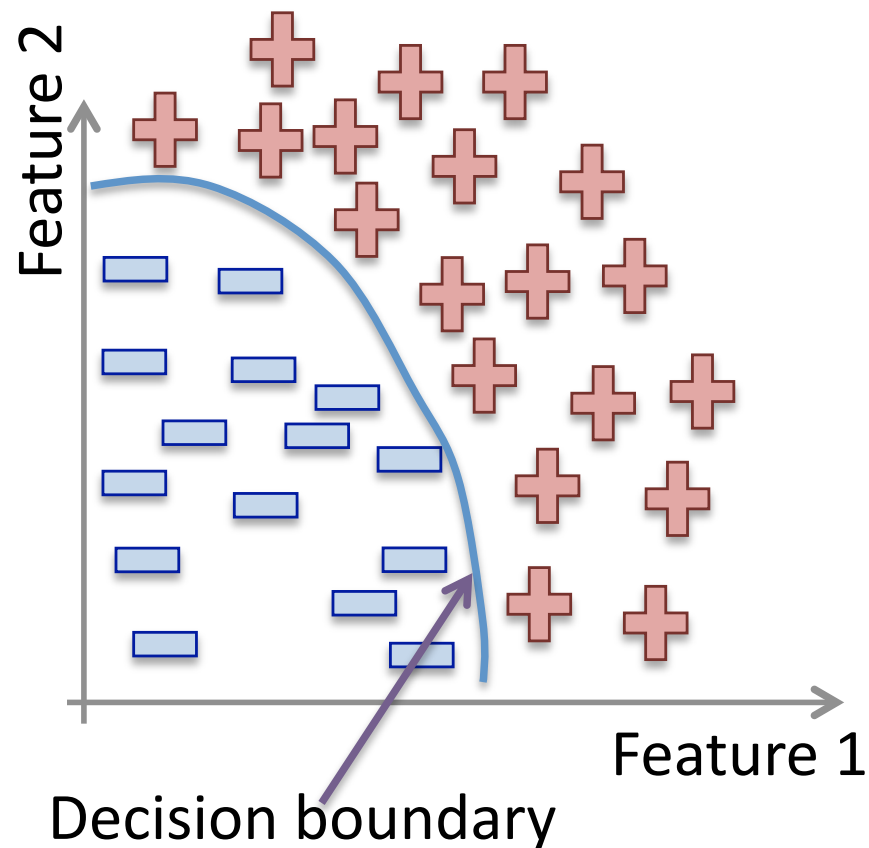
Methods: support vector machines, neural networks, decision trees, k-nearest neighbors, Naive Bayes, etc.

Supervised Learning

Classification

©aiphabet, Inc.
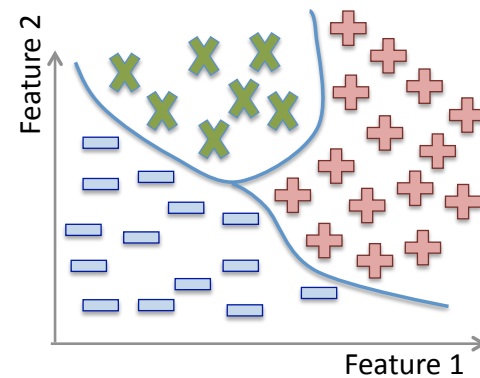
Supervised Learning

Training data: "examples" $\vec{x}$ with "labels" y

$$(\vec{x_1}, y_1), \ldots, (\vec{x_n}, y_n) \; / \; \vec{x_i} \in \mathbb{R}^d$$

- **Regression:** y is a real value, $y \in \mathbb{R}$

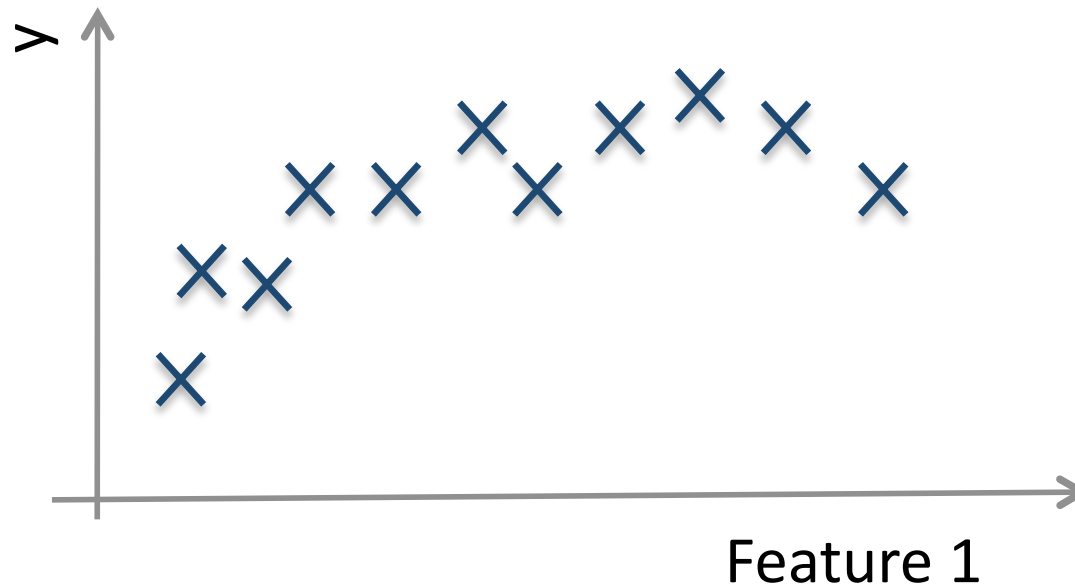$$f : \mathbb{R}^d \longrightarrow \mathbb{R} \qquad\qquad f \text{ is called a } \textbf{regressor}.$$

Example: amount of credit, weight of fruit

Supervised Learning

Regression

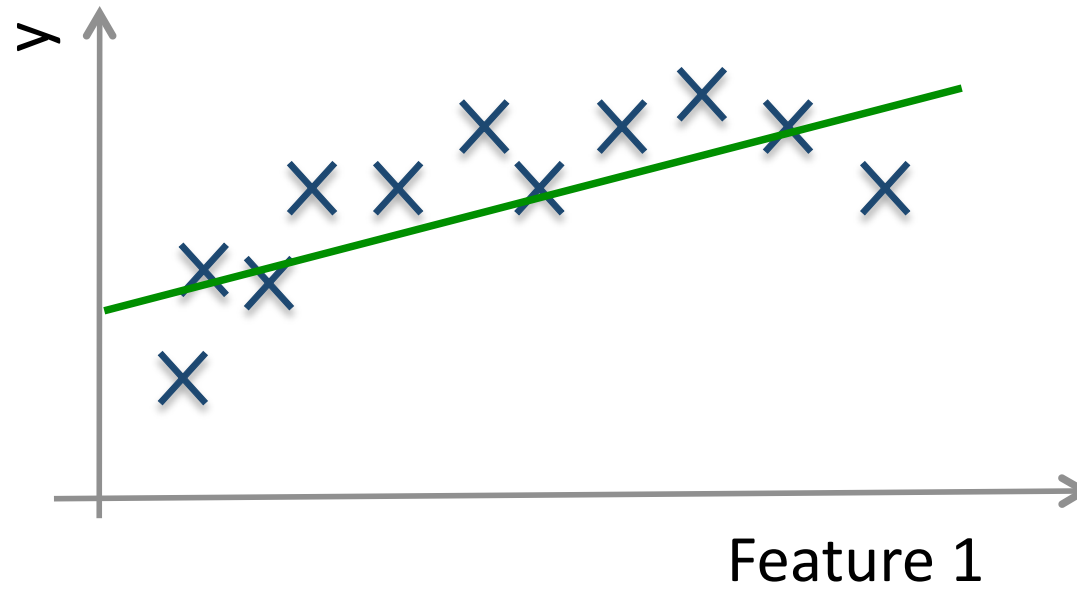

Example: income in function of age, weight of the fruit in function of its length
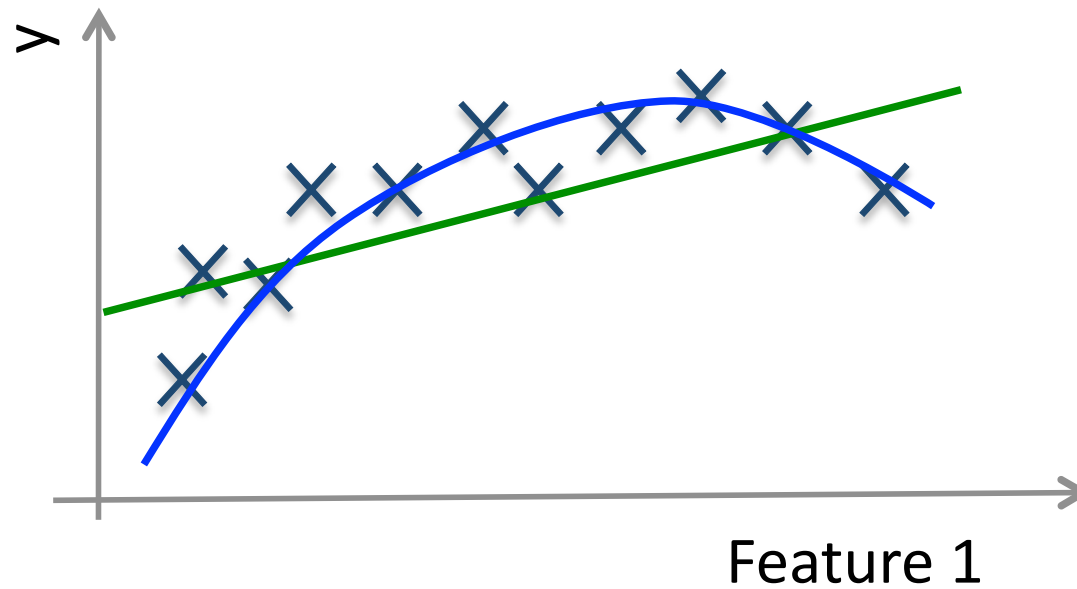
Supervised Learning

Regression:

Supervised Learning

Regression:

Supervised Learning

Regression:

©aiphabet, Inc.