Introduction to Supervised Learning

Kookjin Lee

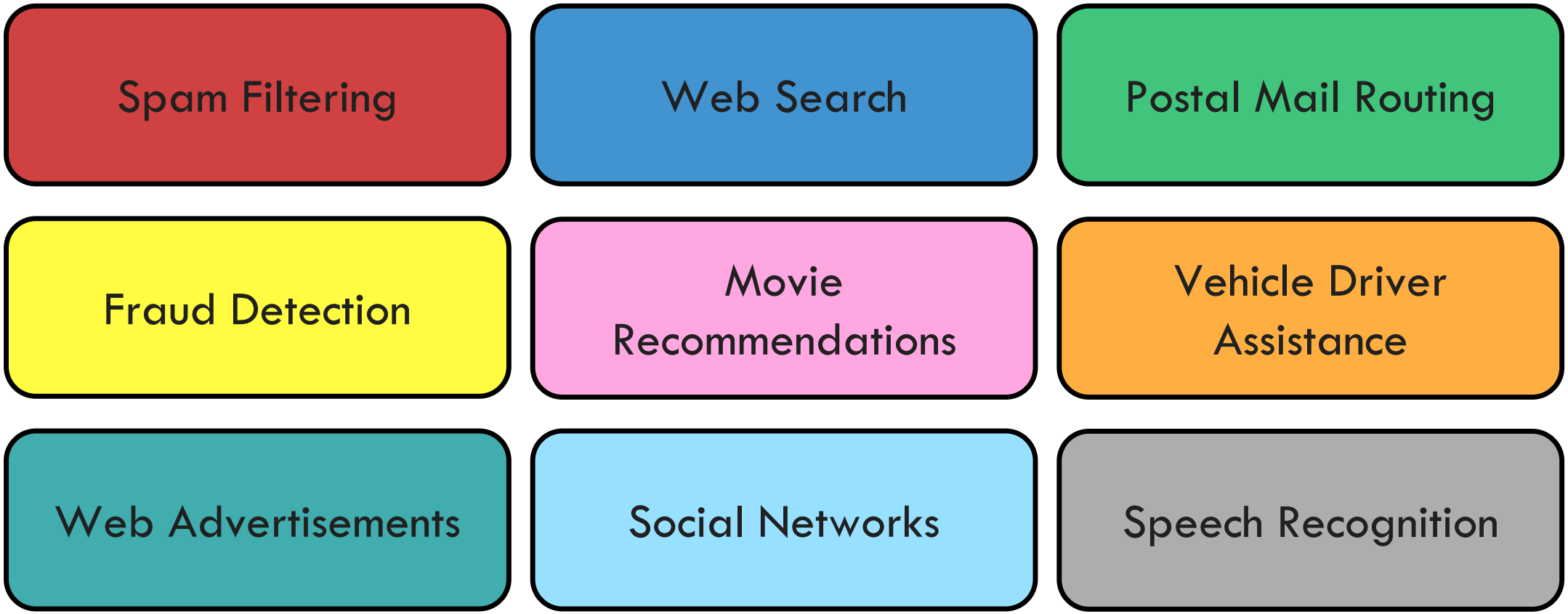
(kookjin.Lee@asu.edu)

Introduction to Machine Learning (ML) - definitions

- Machine Learning is the field of study that gives the computer the ability to learn without being explicitly programmed.
  -- Arthur Samuel (1959)

- A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.
  -- Tom Michell (1998)

Machine Learning in our daily lives



Three modes of Machine Learning

- **Supervised learning**: data with labels $(x^{(i)}, y^{(i)})$

$$\{(x^{(i)}, y^{(i)})\}_{i=1}^{n_{\text{train}}}$$

- Unsuperivsed learning: data w/o labels $x^{(i)}$

$$\{x^{(i)}\}_{i=1}^{n_{\text{train}}}$$

Supervised — data points have known outcome

Unsupervised — data points have unknown outcome

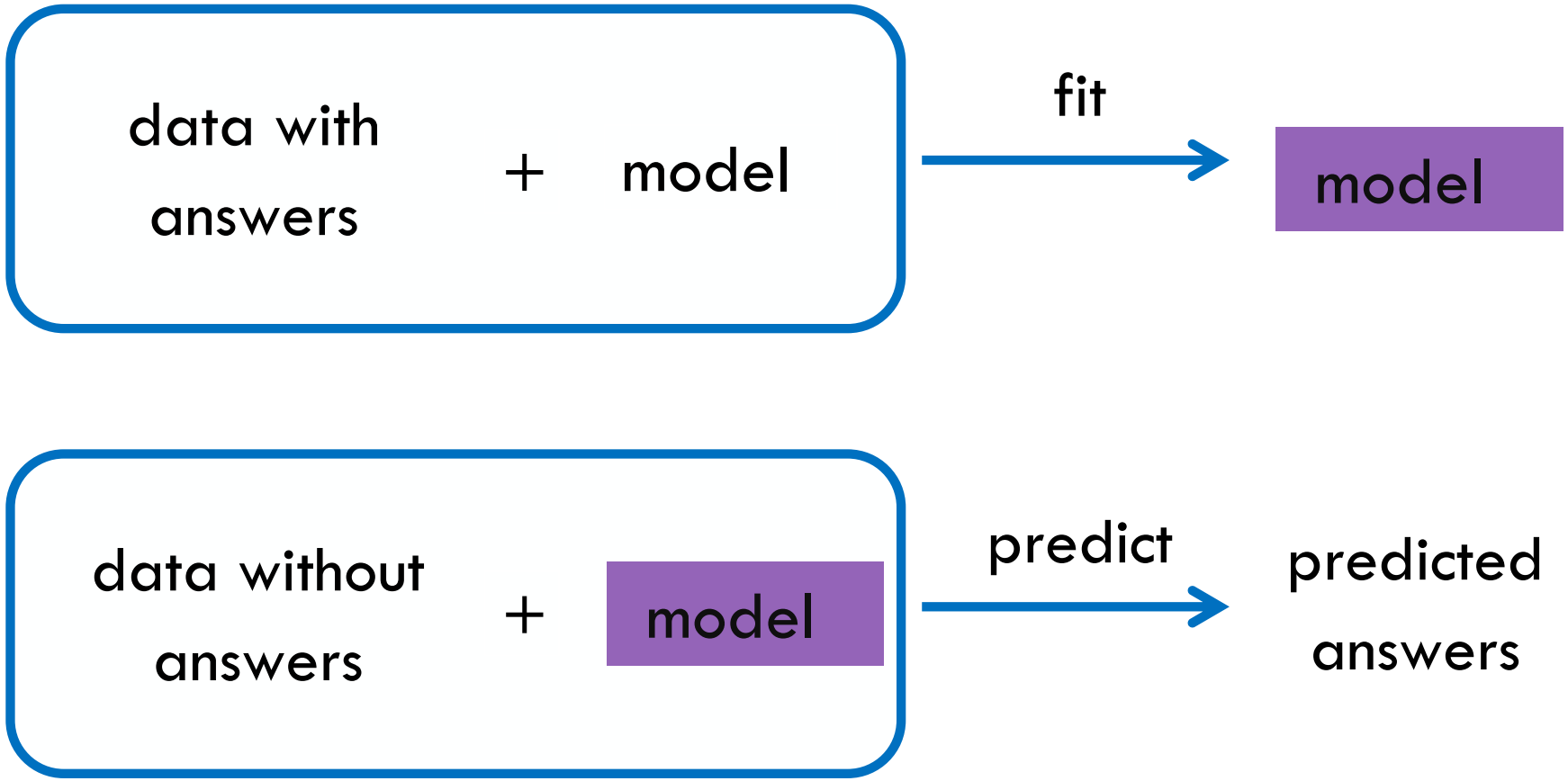
Types of supervised learning

- Two types:
  - Regression: $y^{(i)} \in \mathbb{R}$
  - Classification: $y^{(i)} \in \{0, 1\}, \quad y^{(i)} \in \{\text{'Cat'}, \text{'Dog'}, \text{'Fox'}, \text{'Cow'}\}$

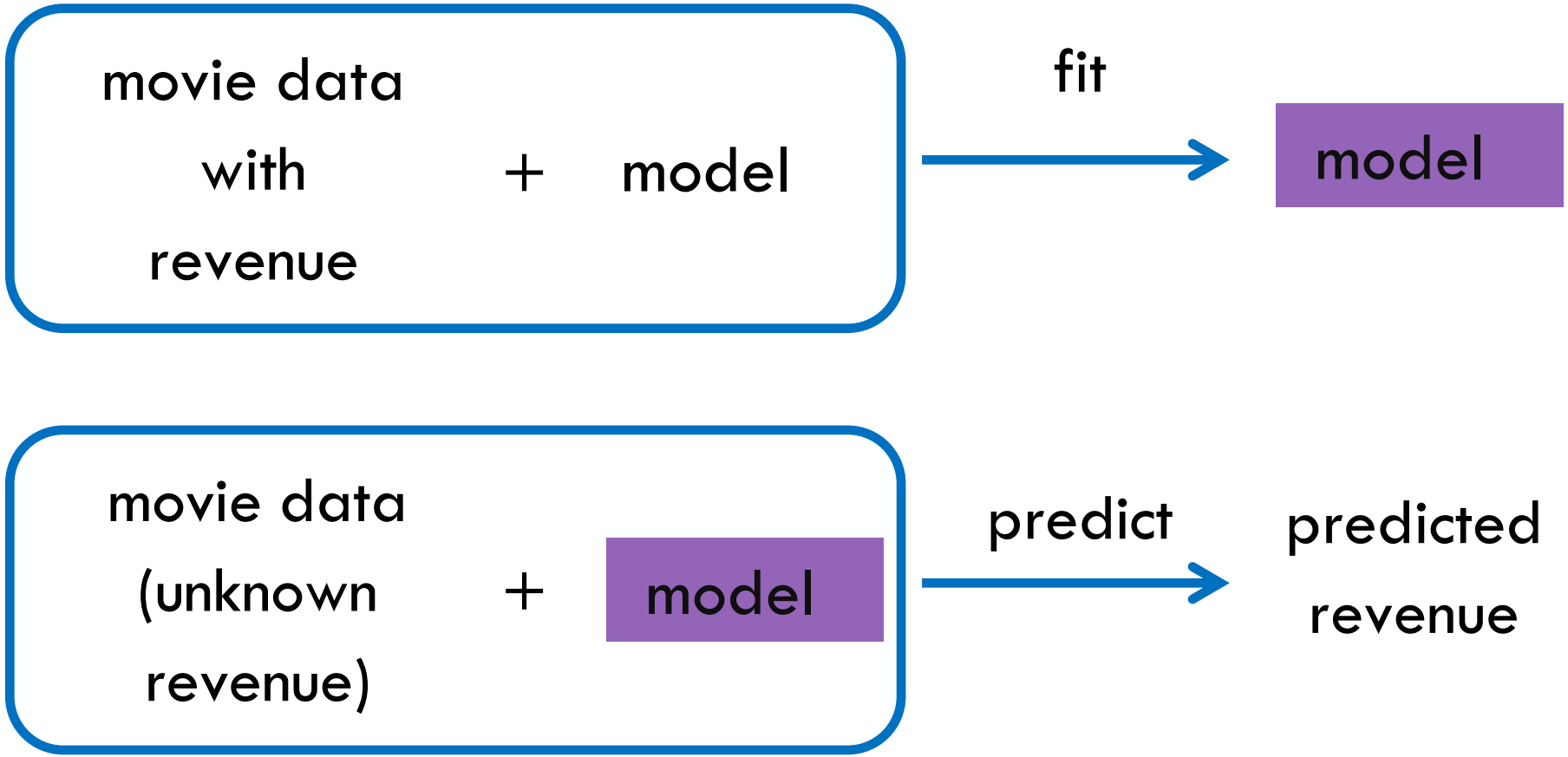
outcome is continuous (numerical)

outcome is a category

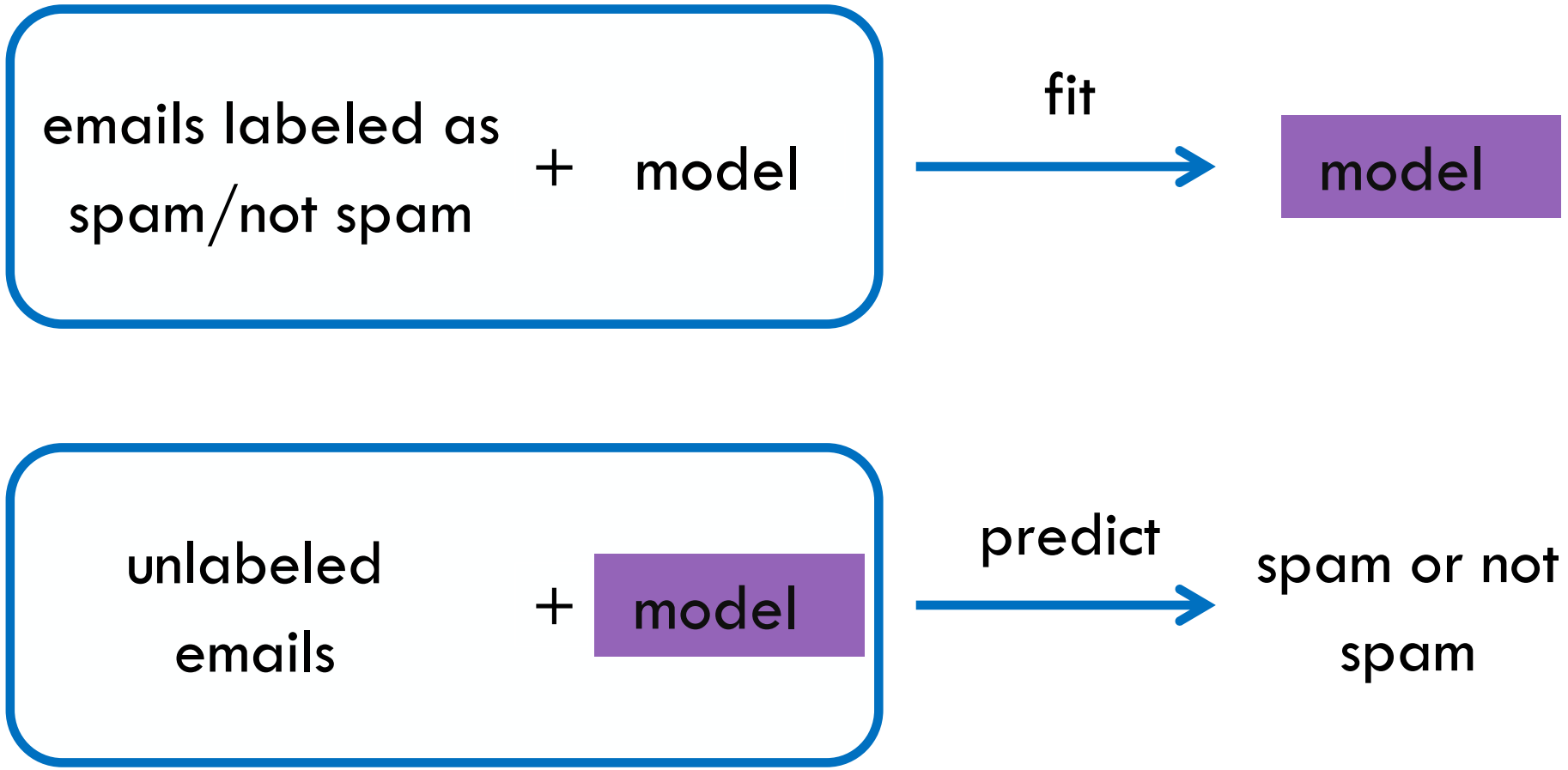
Supervised learning overview



Regression: numeric answers

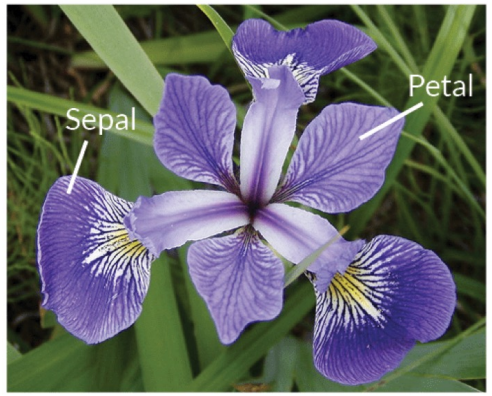


Classification: categorical answers



Notations

- $x^{(i)}$: input variable, input features
- $y^{(i)}$: output variable, target variable
- $\{(x^{(i)}, y^{(i)})\}_{i=1}^{n_{\text{train}}}$: training set



Iris Versicolor

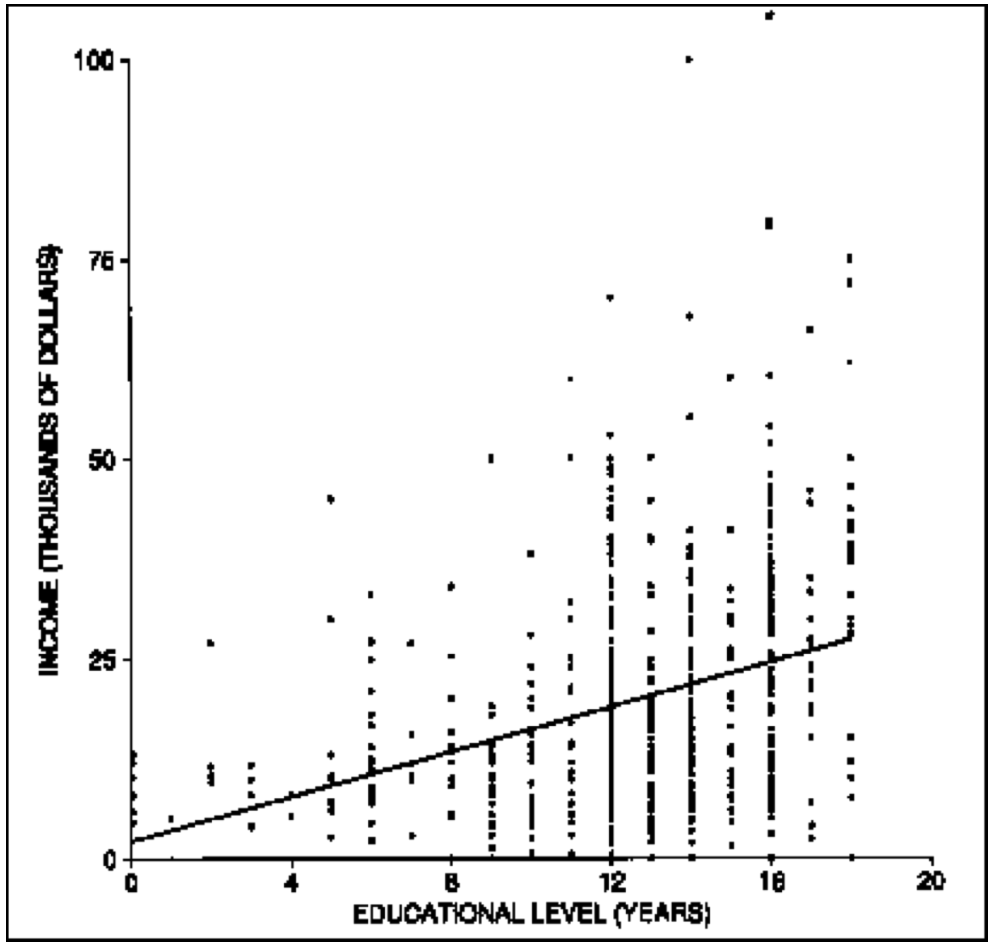
Iris Setosa

Iris Virginica

sepal length	sepal width	petal length	petal width	species
6.7	3.0	5.2	2.3	virginica
6.4	2.8	5.6	2.1	virginica
4.6	3.4	1.4	0.3	setosa
6.9	3.1	4.9	1.5	versicolor
4.4	2.9	1.4	0.2	setosa
4.8	3.0	1.4	0.1	setosa
5.9	3.0	5.1	1.8	virginica
5.4	3.9	1.3	0.4	setosa
4.9	3.0	1.4	0.2	setosa
5.4	3.4	1.7	0.2	setosa

Introduction to Machine Learning (ML) -- examples

- Example 5: regression



Notations

- One more notation:
  - $h_\theta(x)$ or $h(x;\theta)$: hypothesis, (parameterized) model, (parameterized) function

Summary

- Given the training set, $\{(x^{(i)}, y^{(i)})\}_{i=1}^{n_{\text{train}}}$, we want to learn a model (specifically, model parameters) that provides a mapping from $x^{(i)}$ to $y^{(i)}$