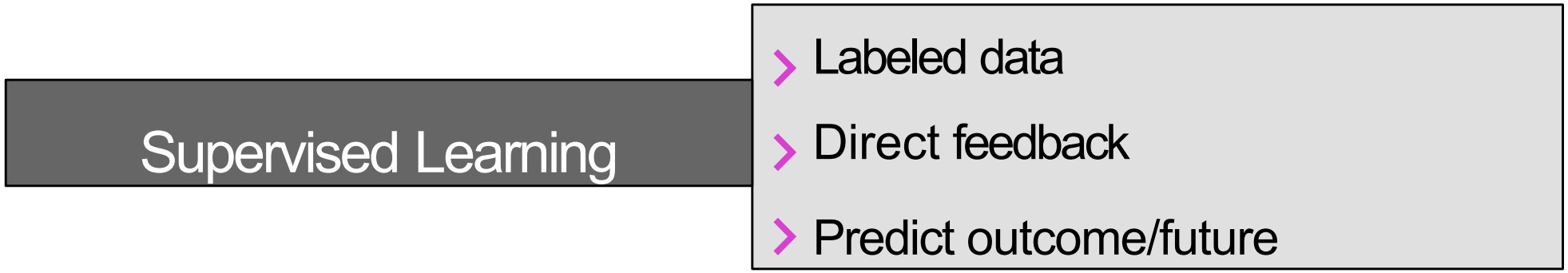
# Machine Learning
# CSCE 5215

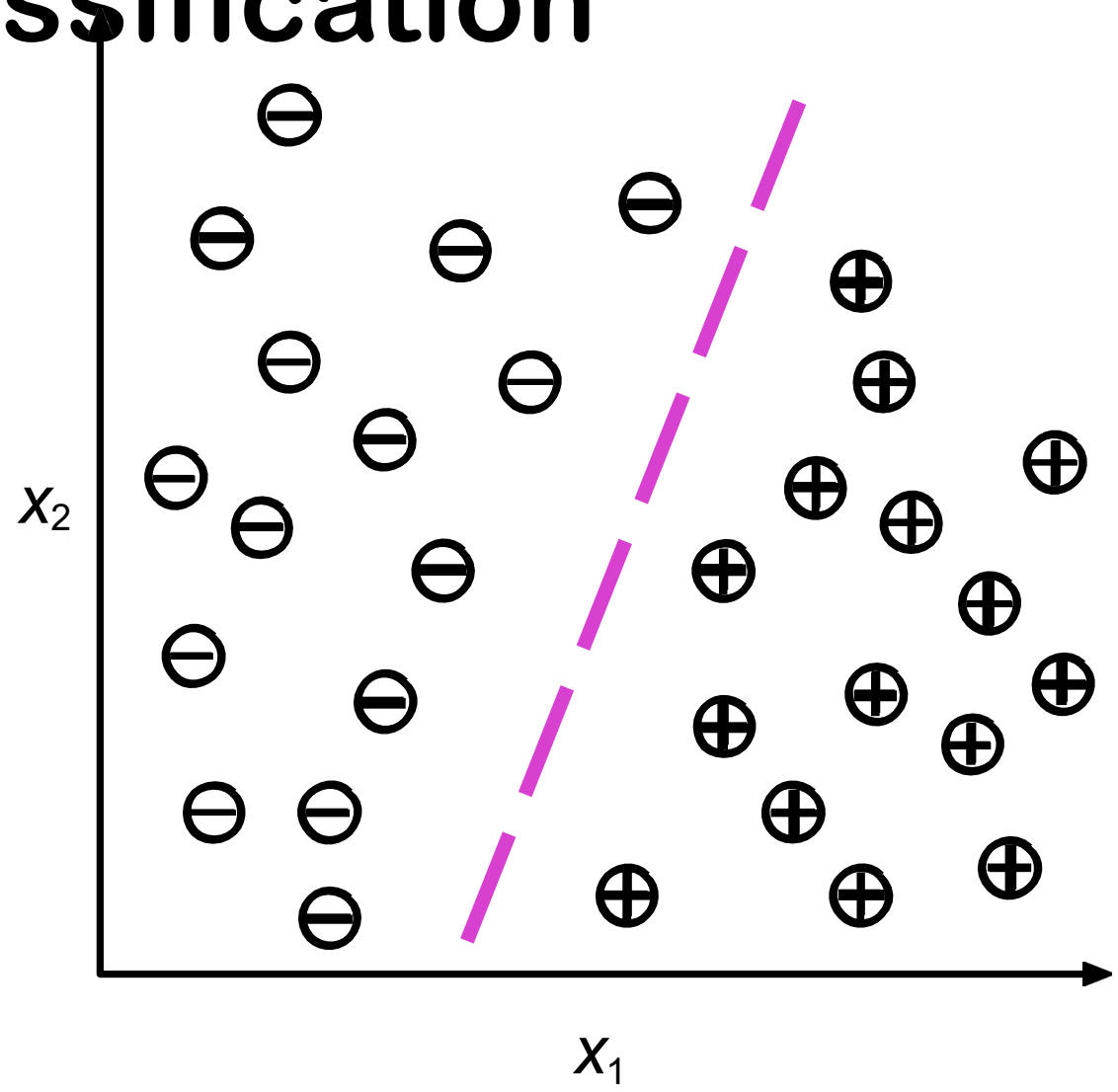
## Categories of Machine Learning
## &
## Data Representation

Instructor: Zeenat Tariq

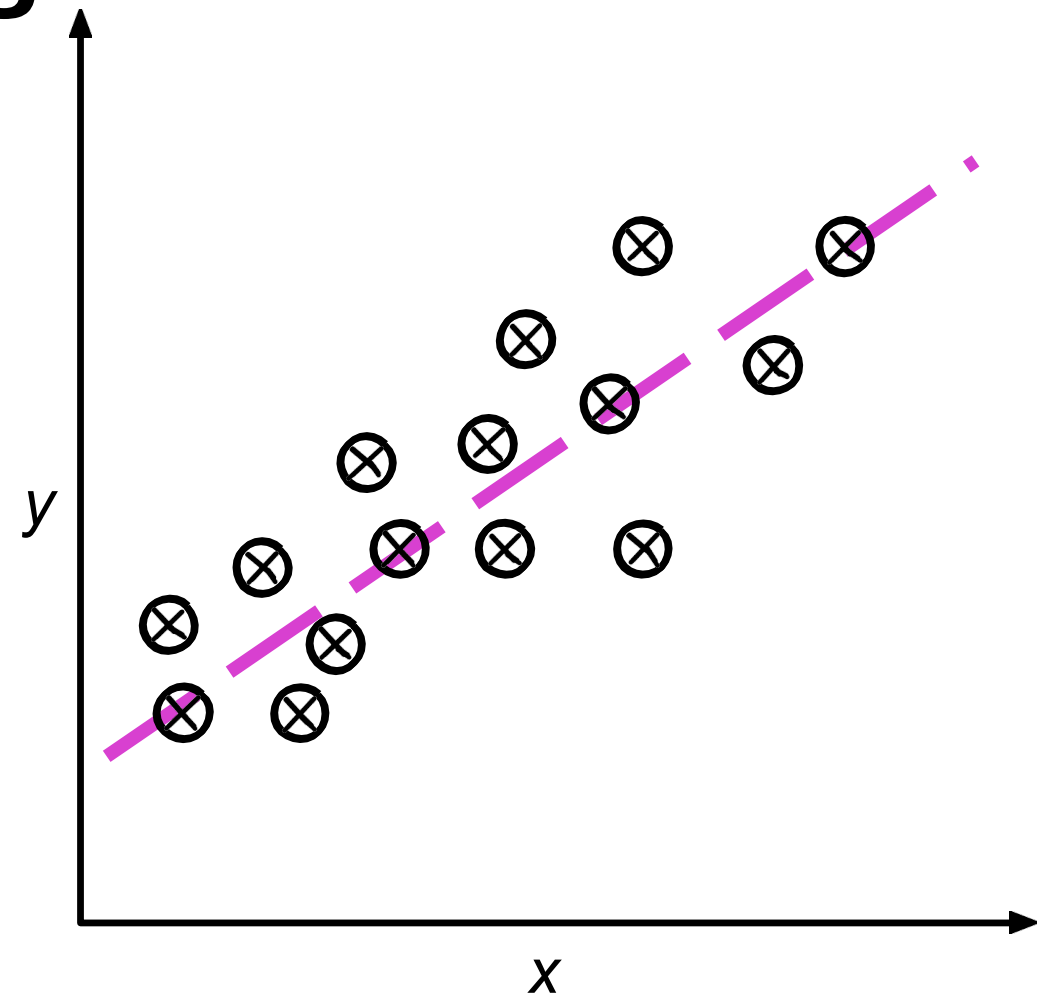
# Categories of Machine Learning



# Supervised Learning: Classification



# Supervised Learning: Regression



# Categories of Machine Learning

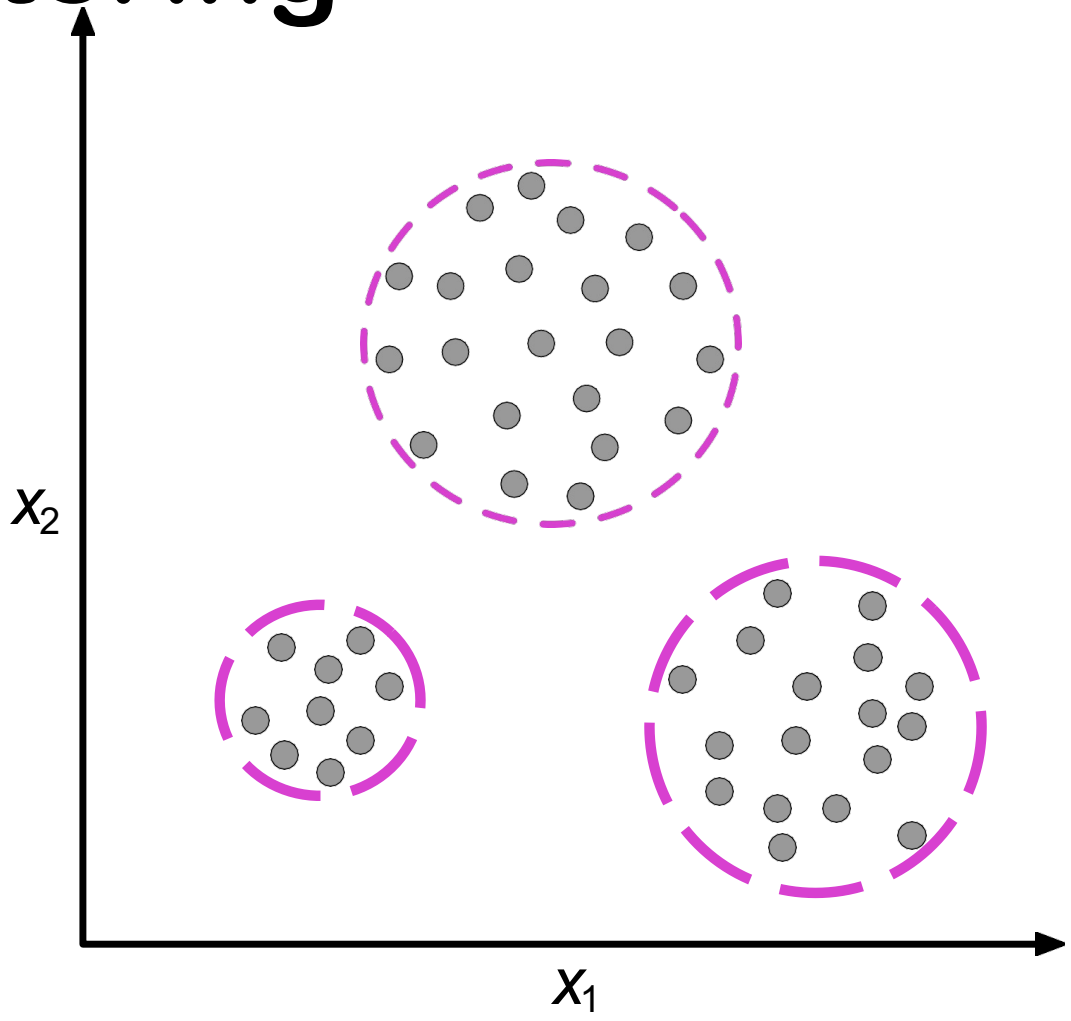
**Supervised Learning**

- Labeled data
- Direct feedback
- Predict outcome/future

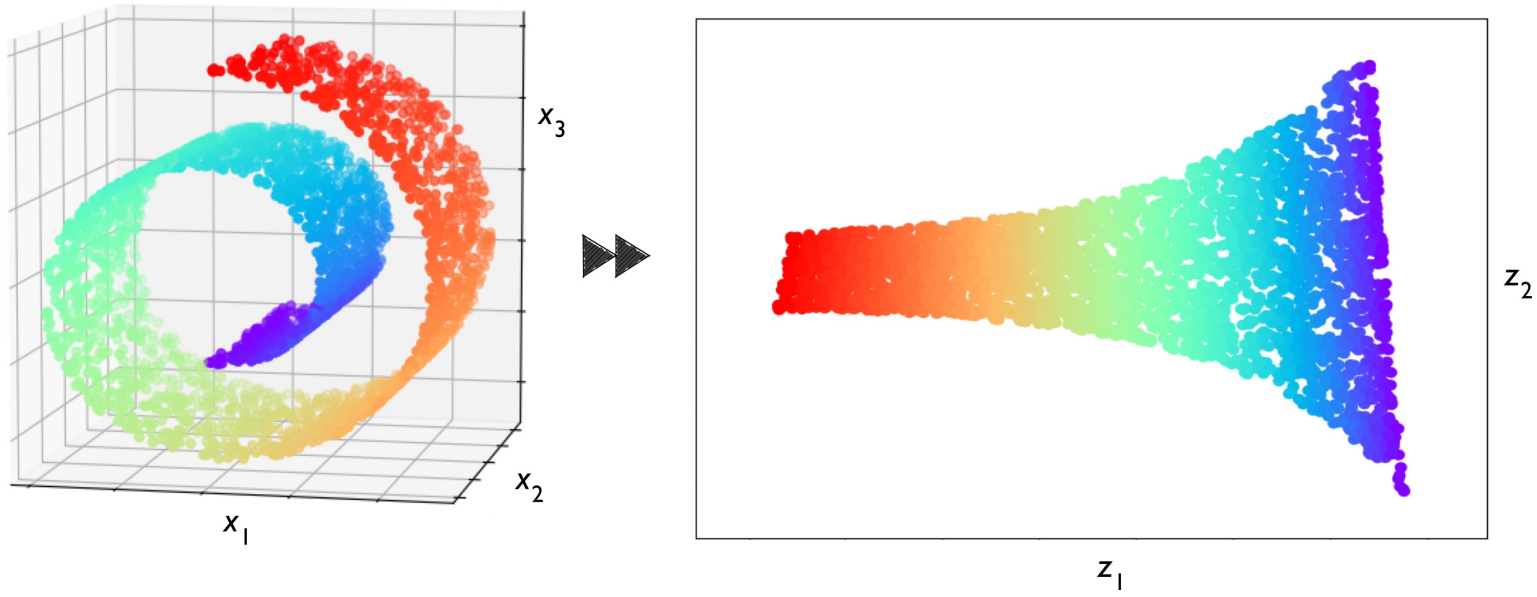
**Unsupervised Learning**

- No labels/targets
- No feedback
- Find hidden structure in data

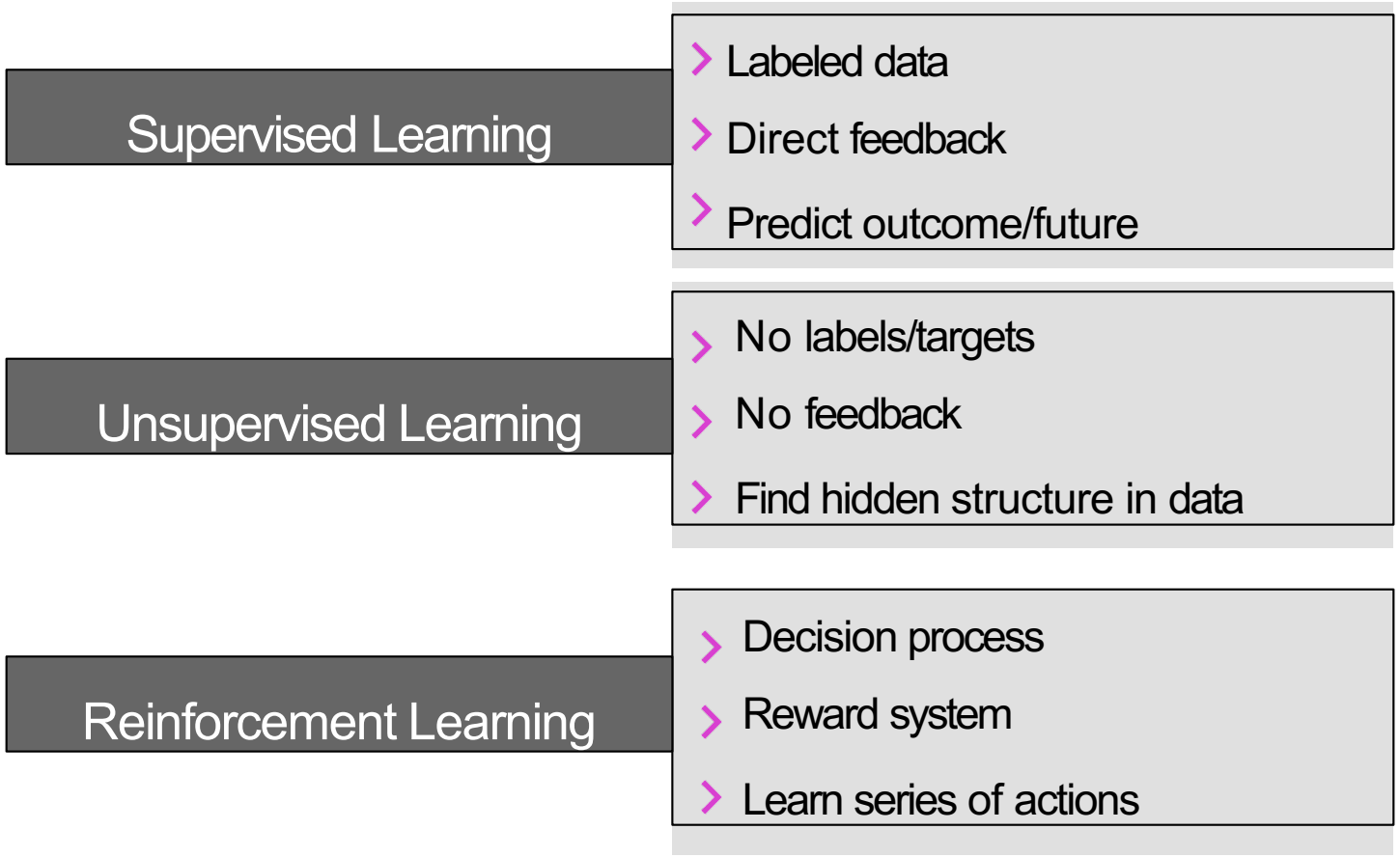
# Unsupervised Learning -- Clustering



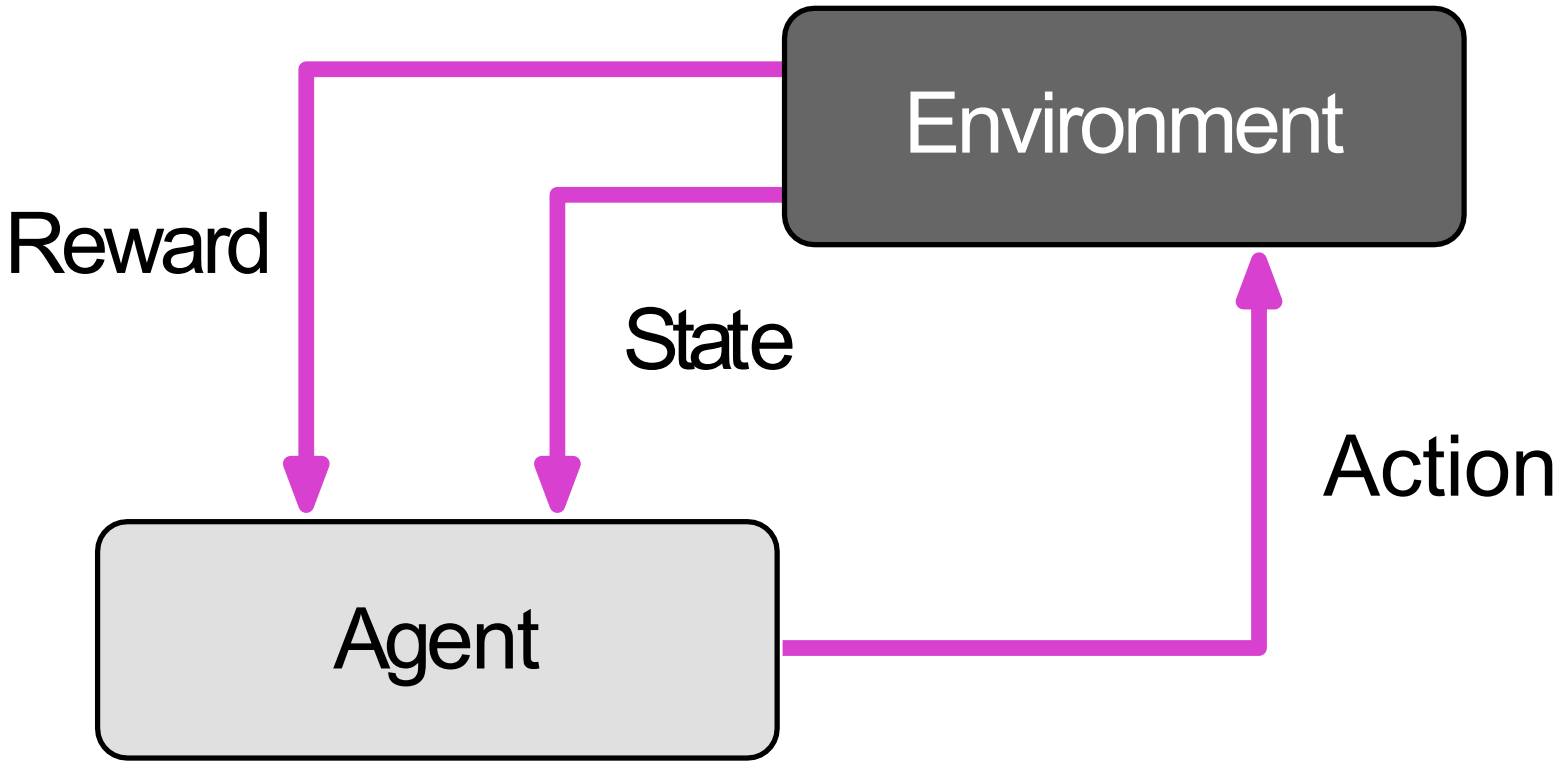
# Unsupervised Learning
# -- Dimensionality Reduction



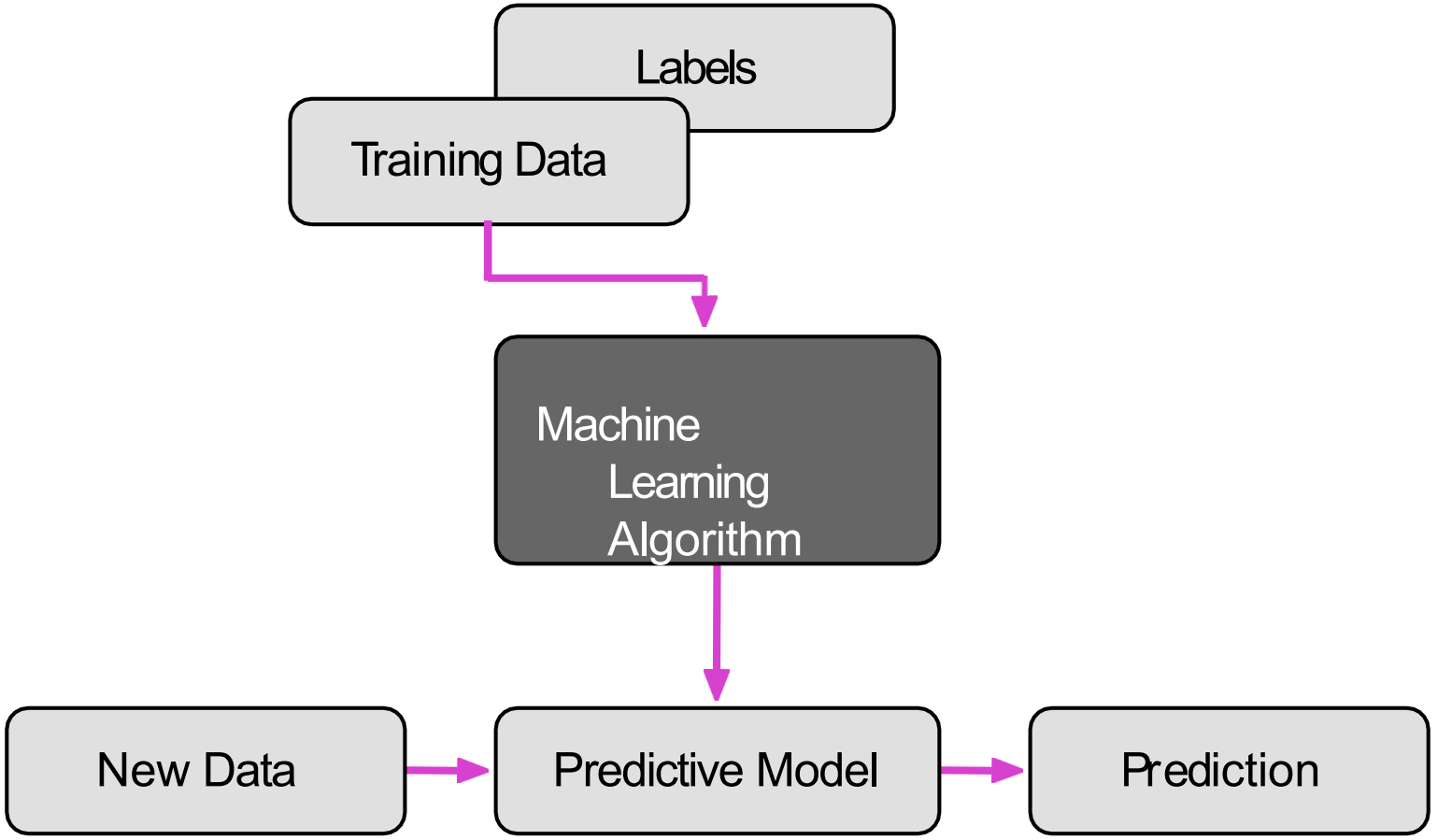
# Categories of Machine Learning



# Reinforcement Learning



# Supervised Learning Workflow -- Overview



# Supervised Learning Notation

Training Set $\mathcal{D} = \{\langle \mathbf{x}^{[i]}, y^{[i]} \rangle, i = 1, \ldots, n\}$,

Unknown function: $f(\mathbf{x}) = y$

Hypothesis: $h(\mathbf{x}) = y$





$h : \mathbb{R}^m \rightarrow \{0, 1)$

$h : \mathbb{R}^m \rightarrow \mathrm{R}$

# Data Representation

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

Feature vector

# Data Representation

Each feature vector corresponds to one training example

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

Feature vector

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_n^T \end{bmatrix}$$

Design Matrix Notation

# Data Representation

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

Feature vector

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_n^T \end{bmatrix}$$

Design Matrix

$$\mathbf{X} = \begin{bmatrix} x_1^{[1]} & x_2^{[1]} & \cdots & x_m^{[1]} \\ x_1^{[2]} & x_2^{[2]} & \cdots & x_m^{[2]} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{[n]} & x_2^{[n]} & \cdots & x_m^{[n]} \end{bmatrix}$$

Design Matrix

# Data Representation

$m = 4$

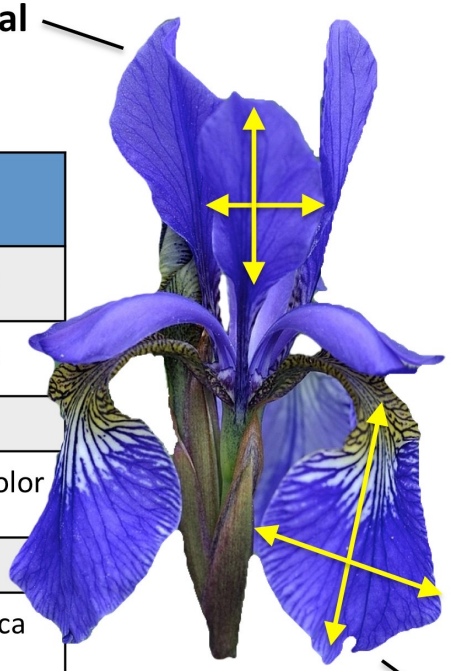
$n = 150$

Samples

| | Sepal length | Sepal width | Petal length | Petal width | |
|---|---|---|---|---|---|
| 1 | 5.1 | 3.5 | 1.4 | 0.2 | Setosa |
| 2 | 4.9 | 3.0 | 1.4 | 0.2 | Setosa |
| ... | | | | | |
| 50 | 6.4 | 3.5 | 4.5 | 1.2 | Versicolor |
| ... | | | | | |
| 150 | 5.9 | 3.0 | 5.0 | 1.8 | Virginica |

Features

Class labels
(targets)



# Data Representation

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} y^{[1]} \\ y^{[2]} \\ \vdots \\ y^{[n]} \end{bmatrix}$$

Input features