



## Supervised Learning

### DATA TYPE

Labeled data ( $X, y$ )

### GOAL

Predict  $y$  from  $X$

- ✓ Requires labeled data
- ✓ High accuracy
- ✓ Expensive labeling
- ✓ Domain expertise needed

Labeling Cost



## Unsupervised Learning

### DATA TYPE

Only  $X$  (unlabeled)

### GOAL

Find structure in  $X$

- ✓ No labels needed
- ✓ Discover patterns
- ✓ Highly scalable
- ✓ Cost-effective

Labeling Cost



## Semi-supervised Learning

### DATA TYPE

Small labeled + Large unlabeled

### GOAL

Leverage both data types

- ✓ Best of both worlds
- ✓ Balanced approach
- ✓ Practical solution
- ✓ Often best results

Labeling Cost