Introduction to Machine Learning

MODULE 2 Categories of Machine Learning

Modules for this course

1. Overview: What is Machine learning
2. **Categories of machine learning**
3. Notation
4. Machine Learning application approach
5. Recommender Systems
6. Building a Recommender Engine

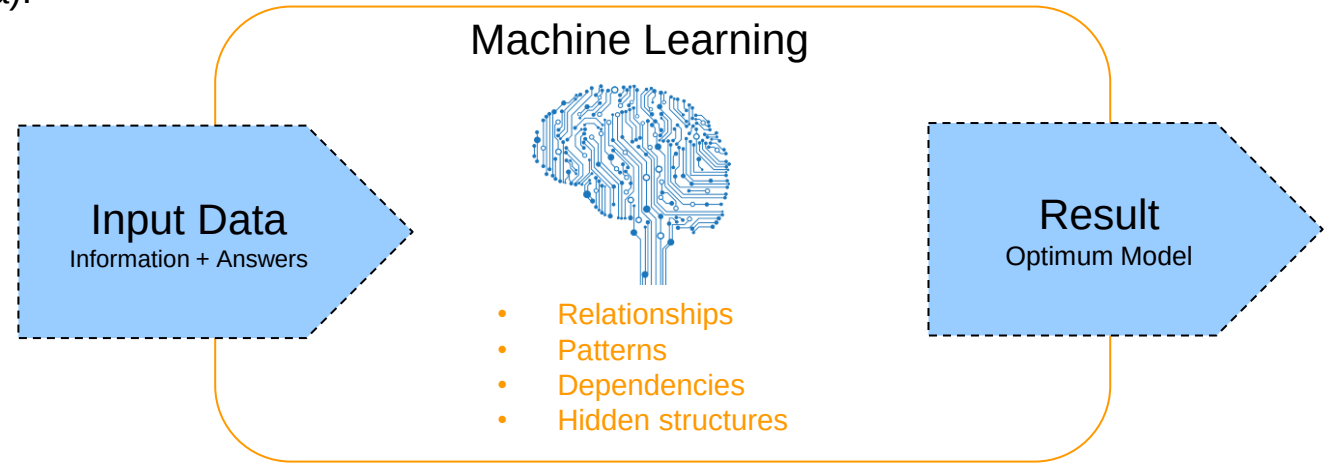
Module 2

Categories of Machine Learning

Machine Learning – Breaking it down

Supervised and Unsupervised Learning

1. **Supervised learning** - we already know the answers we want (found in past or completed data).



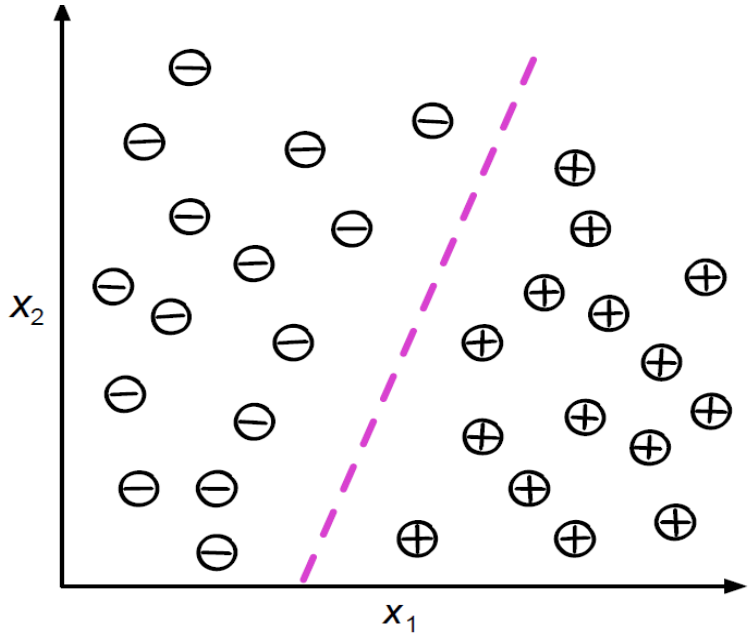
2. **Unsupervised learning** - we want to find unknown structures or trends.
3. **Reinforcement learning** – we use trial and error with rewards

Supervised Learning: Classification

Labeled Data

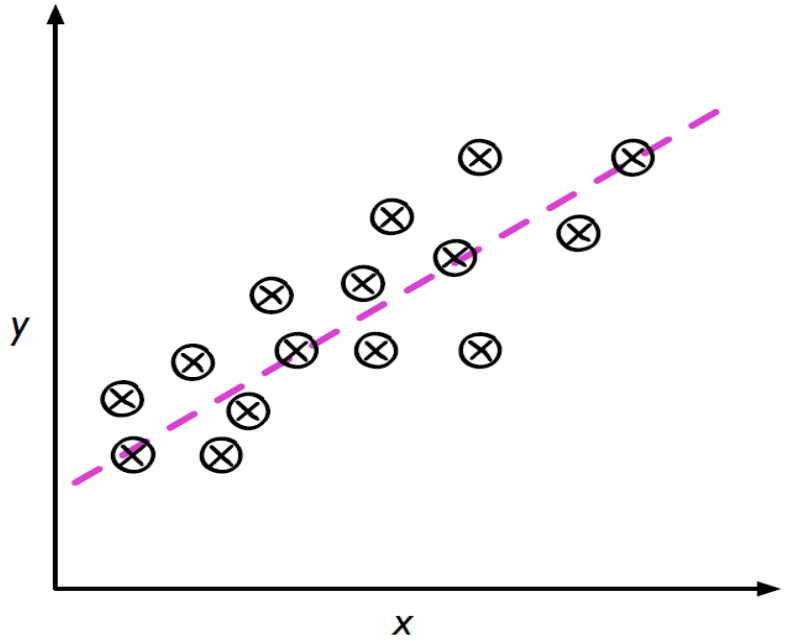
Direct Feedback

Predict Outcome/Future



Supervised Learning: Regression

Labeled Data | Direct Feedback | Predict Outcome/Future

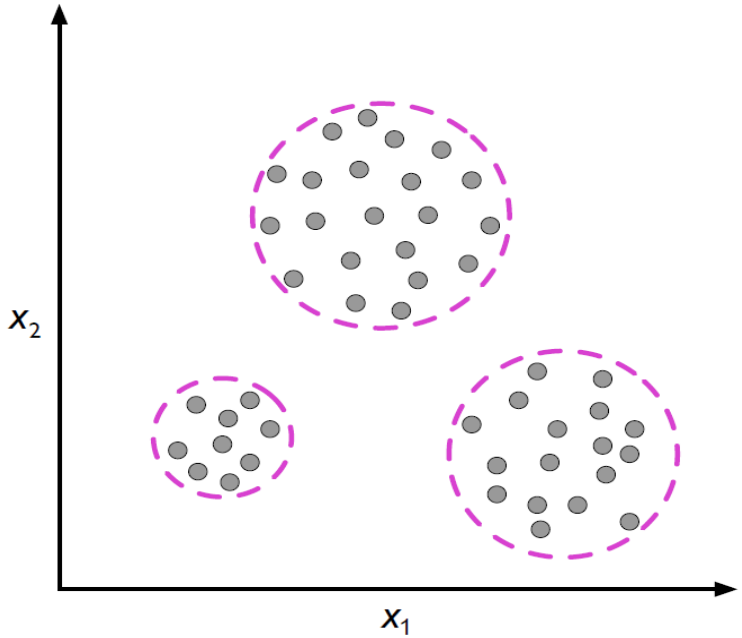


Unsupervised Learning: Clustering

No Labels/targets

No Feedback

Find Hidden patterns

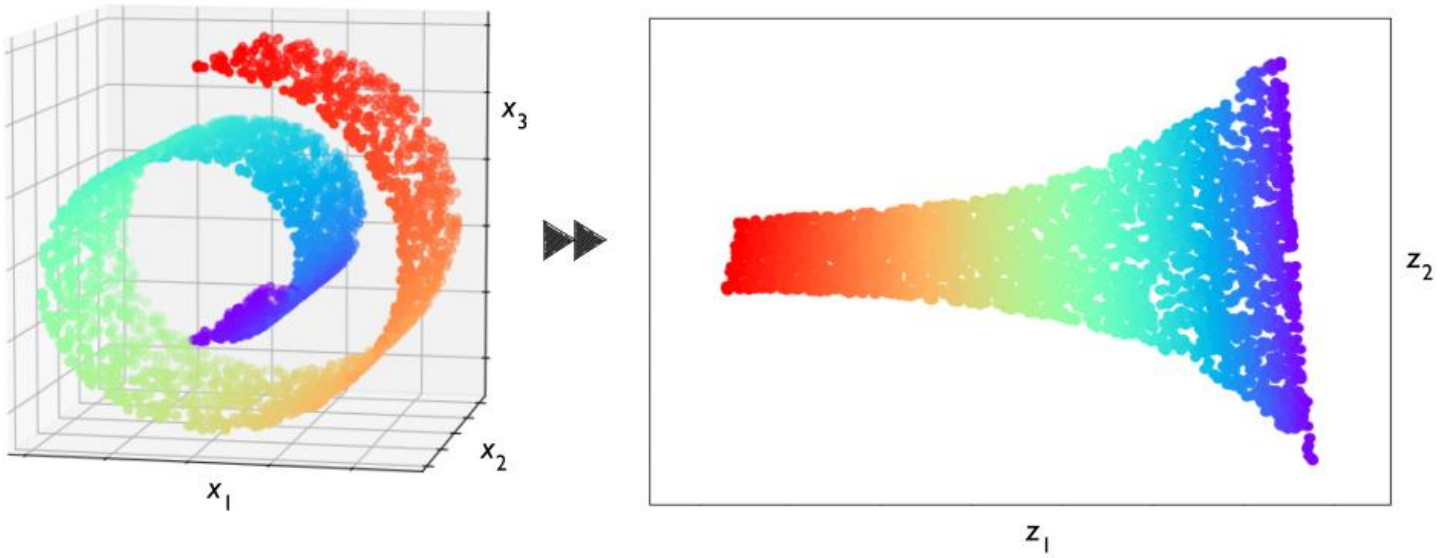


Unsupervised Learning: Dimensionality Reduction

No Labels/targets

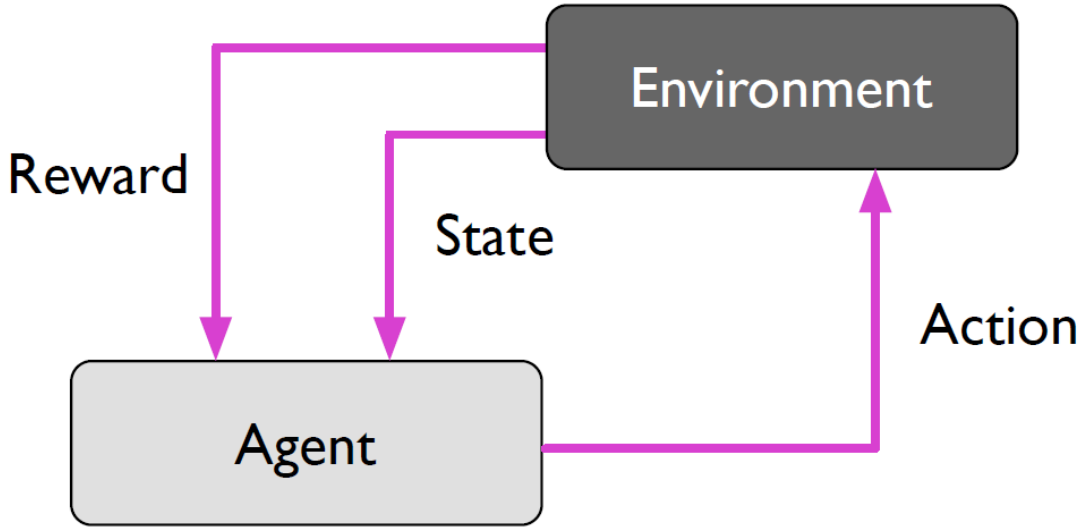
No Feedback

Find Hidden patterns

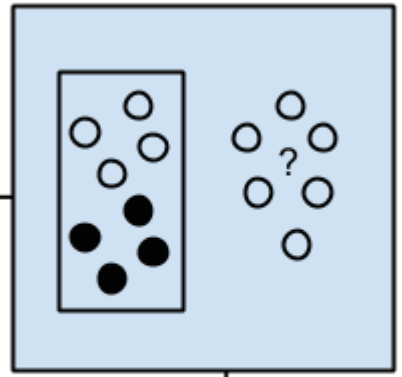


Reinforcement Learning:

Decision Process | Reward System | Learn Series of Actions



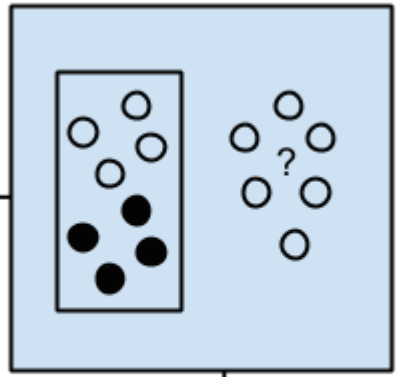
Supervised Learning – Simple Example



Math Quiz #1 - Teacher's Answer Key

1)	2	4	5	=	3	5)	6	2	2	=	10
2)	5	2	8	=	2	6)	3	1	1	=	2
3)	2	2	1	=	3	7)	5	3	4	=	11
4)	4	2	2	=	6	8)	1	8	1	=	7

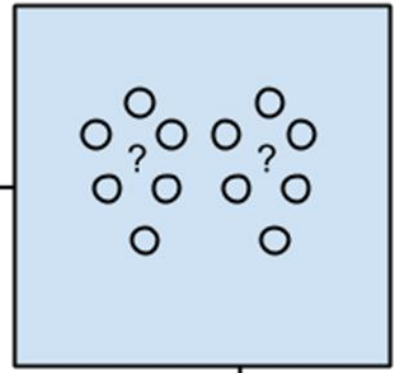
Supervised Learning – Simple Example



Math Quiz #1 - Teacher's Answer Key

1) 2 x 4 - 5 = 3
2) 5 x 2 - 8 = 2
3) 2 x 2 - 1 = 3
4) 4 x 2 - 2 = 6
5) 6 x 2 - 2 = 10
6) 3 x 1 - 1 = 2
7) 5 x 3 - 4 = 11
8) 1 x 8 - 1 = 7

Unsupervised Learning – Simple Example



Math Quiz #1 - Teacher's Answer Key

1) 2 4 5 =
2) 5 2 8 =
3) 2 2 1 =
4) 4 2 2 =
5) 6 2 2 =
6) 3 1 1 =
7) 5 3 4 =
8) 1 8 1 =

Supervised & Unsupervised Learning

Supervised Learning:

Predicting values. **Known** targets.
User inputs correct answers to learn from. Machine uses the information to guess new answers.

REGRESSION:
Estimate continuous values
(Real-valued output)

CLASSIFICATION:
Identify a unique class
(Discrete values, Boolean, Categories)

Unsupervised Learning:

Search for structure in data. Unknown targets.
User inputs data with undefined answers. Machine finds useful information hidden in data

CLUSTER ANALYSIS:
Group into sets

DENSITY ESTIMATION:
Approximate distribution

DENSITY REDUCTION:
Select relevant variables

Supervised & Unsupervised Learning

Supervised Learning:

Classification

- Decision Trees
- K-Nearest Neighbors
- Support Vector Machine
- Logistic Regression
- Naïve Bayes
- Random Forests

Regression

- Linear Regression
- Ordinary Least Squares Regression
- LOESS (Local Regression)
- Neural Networks

Unsupervised Learning:

Cluster Analysis

- K-Means Clustering
- Hierarchical Clustering

Dimension Reduction

- Principal Component Analysis (PCA)
- Linear Discriminant Analysis (LDA)

Exercise

Building classification models is one of the most important data science use cases. Classification models are models that predict a categorical label. A few examples of this include predicting whether a customer will churn or whether a bank loan will default. In this guide, you will learn how to build and evaluate a classification model in R. We will train the logistic regression algorithm, which is one of the oldest yet most powerful classification algorithms.

Summary

- We now have a good understanding of the categories of machine learning
- We can differentiate between supervised, unsupervised and reinforcement learning
- We are now ready to do a simple exercise of supervised learning.