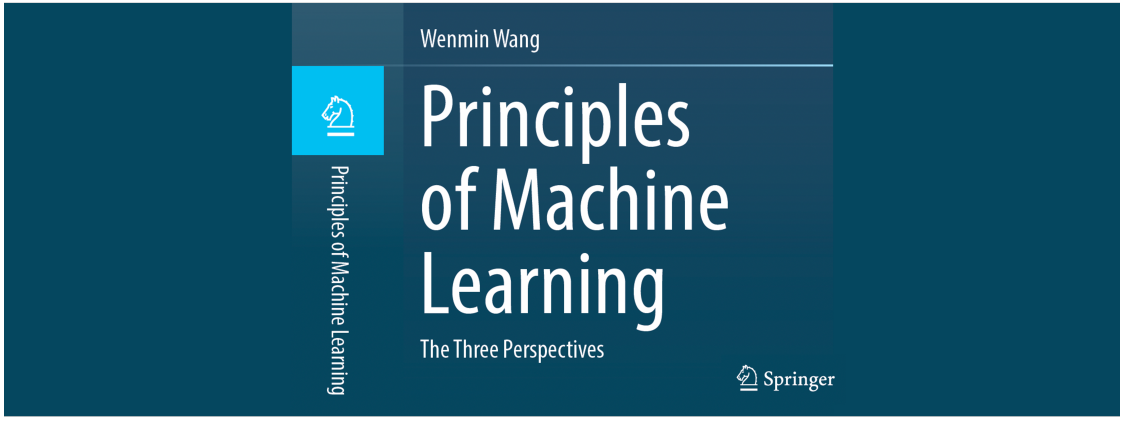

Wenmin Wang
Principles of Machine Learning
The Three Perspectives
Principles of Machine Learning
Springer

Part I Perspectives

2 On Perspectives

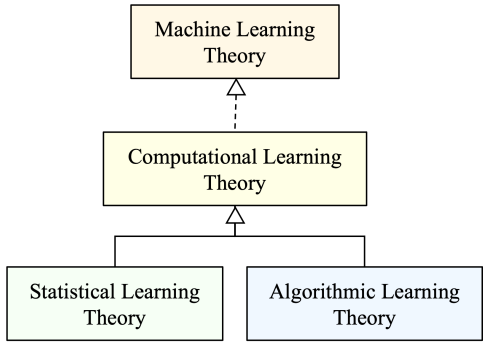
Contents

Learning and Learning Theory

- Learning types:
  Narrow learning: acquiring knowledge or skills through means such as textbooks;
  Broad learning: gaining experience or skills in life.
- Learning styles:
  A series of theories aimed at explaining individual learning differences.
- Learning machines:
  Proposed by Alan Turing in 1950.
- Machine learning:
  Proposed by Arthur Samuel in 1959.
- Learning Theory:
  A topic studied in both education and machine learning.

Machine Learning Theories

- Computational learning theory:
  a formalized learning theory based on probability theory.
- Statistical learning theory:
  a machine learning framework based on statistics.
- Algorithmic learning theory:
  a mathematical framework for analyzing machine learning problems and algorithms.



Machine Learning Theories

- What are perspectives
  - Related to the angle and dimension of looking at a problem.
  - One angle can only see one side, and one dimension can only see one line.
  - Only from multiple angles and dimensions can we see full picture of things.
- For machine learning
  - Studying it from different perspectives can bring it closer to its essence.
- Three perspectives studying machine learning
  - Theoretical perspective.
  - Methodological perspective.
  - Practical perspective.

Major Foundations

Mathematics	Probability Theory
	Statistics
	Linear Algebra
	Calculus
	Geometry
	Functional Analysis
Logic	
Information Theory	
Decision Theory	
Neuroscience	

Mathematics: Probability Theory

What is Probability Theory

- Be a mathematical theory about random phenomena and uncertain problems.
- Be used to study the inherent laws in these phenomena and problems and to analyze the possibility of various outcomes using mathematical methods.

Two Schools

- Frequentist Probability Theory
- Bayesian Probability Theory

Theorems

- Law of large numbers
- Central limit theorem
- Bayes' theorem.

Mathematics: Early Representatives of Probability Theory

Girolamo Cardano (1501-1576)
Italian Polymath
Book on Games of Chance

Pierre de Fermat (1607–1665)
French Mathematician

Blaise Pascal (1623–1662)
French Mathematician

Problem of Points in gambling

Christiaan Huygens (1629-1695)
Dutch Mathematician
On Reasoning in Games of Chance

Jacob Bernoulli (1654-1705)
Swiss Mathematician
Law of large numbers (LLN)

Thomas Bayes (1701-1761)
British Mathematician
Bayes' theorem

Pierre-Simon Laplace (1749-1827)
French Polymath
Classical definition of probability

Aleksandr Lyapunov (1857-1918)
Russian Mathematician
Central limit theorem (CLT)

Mathematics: Statistics

What is Statistics

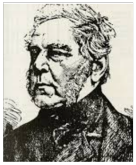
- Be a discipline about the collection, organization, representation, analysis, interpretation, and prediction based on data.

What is Mathematical Statistics

- Be the application of probability theory in statistics, as opposed to techniques for collecting statistical data.

Mathematics: Early Representatives of Statistics

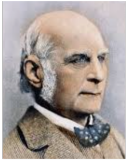
Girolamo Ghilini (1589-1668)
Italian Scholar
Proposed the term statistics



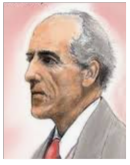
John Graunt (1620-1674)
British demographer

William Petty (1623-1687)
British demographer

Census and statistical methods



Francis Galton (1822-1911)
British polymath



Karl Pearson (1857-1936)
British mathematician

1st stage: Turned statistics into mathematics

William Gosset (1876-1937)
British Statistician

Ronald Fisher (1890-1962)
British Polymath

2nd stage: Foundation of modern statistics

Egon Pearson (1895-1980)
British Statistician

Jerzy Neyman (1894-1981)
Polish Mathematician

3rd stage: Developement of modern statistics

Mathematics: Linear Algebra, and Calculus

What is Linear Algebra

- It studies systems of linear equations and their linear mappings.
- It involves the operations of scalars, vectors, matrices, and even tensors.
- A kind of continuous mathematics but not discrete mathematics.

Two major branches of Calculus

- Differential calculus:
  it concerns the rates of change, and slopes of curves.
- Integral calculus:
  it can describe displacement, area, volume.

Mathematics: Geometry, and Functional Analysis

What is Geometry

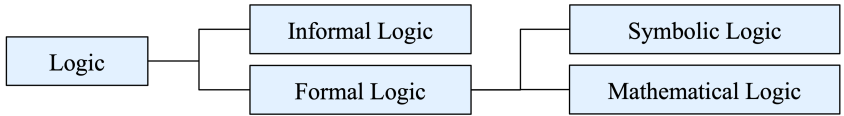
- It studies shapes, sizes, relative positions of figures, and spatial properties.
- Its basic concepts include points, lines, surfaces, angles, distances, and manifolds.
- Geometric deep learning for structured data with geometric structure: graphs, point clouds, grids, geodesics, and manifolds, in Euclidean and non-Euclidean geometry.

What is Functional Analysis

- It studies transformation of functions and their algebraic and topological properties.

Logic

- Systematic study the form of reasoning used to draw correct conclusions.
- Broadly, logic is the analysis and evaluation of arguments.



Symbolic Logic (Classical Logic)	Propositional Logic (Zeroth Order Logic), First Order Logic
Mathematical Logic (Non-Classical Logic)	Modal Logic, Description Logic, Computational Logic, Fuzzy Logic

Information Theory

- A research field for the quantification, storage, and communication of digital information.
- A key measure is entropy.
  - It quantifies the amount of uncertainty involved in the value of a random variable or the outcome of a random process.
  - Joint entropy
  - Conditional entropy
- Mutual information
- Kullback–Leibler divergence (KL divergence)

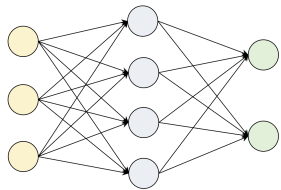
Decision Theory

- A theory of studying the agent's choice.
- A formal framework for choosing the optimal or near-optimal behavior in an uncertain environment.
- Its foundational disciplines:
  - Probability theory.
  - Utility theory.
- Multi-Stage Decisions in ML:
  - Decision tree, Random decision forest, etc.
- Sequential Decision in ML:
  - Reinforcement learning.

Neuroscience

- The science that studies nervous system (brain, spinal cord, and peripheral nerves) and its functions.
  - The neuron is the basic unit of the brain's nervous system.
  - The nervous system to study how the brain processes information.
  - Brain science is one of the main research contents of neuroscience.

- Connectionism establishes its computational model by simulating the function of biological neural networks.
- Artificial neural networks can be seen as an artificial manifestation of the brain.



What the Difficulties We Faced

- There are too many machine learning algorithms that is too hard to count. Moreover, numerous machine learning algorithms emerge every year, every month, and even every day.
- After we studied several machine learning algorithms, we may find that we still haven't gotten out of the maze of machine learning.
- Because we are still unclear about the internal relationship of these algorithms, unable to grasp the overall picture of machine learning.

Different Problems, Different Methods, and Different Theories

Different Problems

- Machine learning is to learn from data or the environment.
- As for data:
  there are different modalities, such as text, image, audio, or video..
- As for environment:
  it is necessary to know the current state and rewards.
- Different modalities of data or feedback of environments, present a diversity of problems.
- So, we need to collate and systematize these problems to understand, i.e.,

What to do.

Different Methods

- The different methods refer to different machine learning algorithms.
- Some algorithms can use a large amount of labeled data to train the machine learning model in order to make predictions on the actual unknown data.
- Some algorithms do not have labeled data but only need to focus on discovering similarities in the data.
- Some algorithms decide the next action based on the current state and reward from the environment.
- So, it is necessary to analyze what type those algorithms belong to, i.e.,

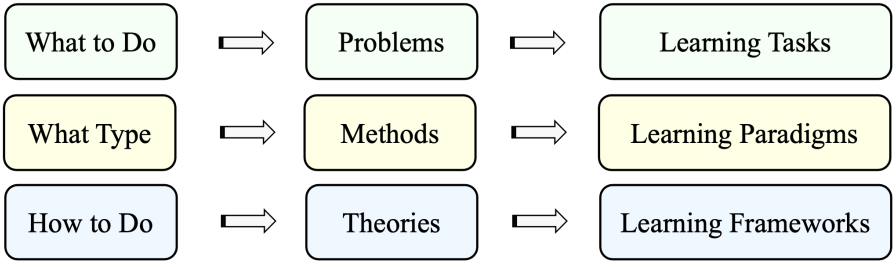
What type.

Different Theories

- Even for the same problem and the same method, there are many machine learning algorithms to consider or choose from.
- Foundational disciplines of machine learning have been introduced, which include mathematics, logic, information theory, decision theory, and neuroscience.
- With these foundational disciplines as the framework, many machine learning algorithms and models have been formed. They should be collated and systematized.
- So, it should analyze the fundamental frameworks of these algorithms, i.e.,

How to do.

Chain of Thought of the Viewpoints



Viewpoints of the three perspectives.

What to Do

- It involves:
  - what “problems” to solve with machine learning,
  - what kind of processing to do on the input data,
  - what the output form is,
  - what action needs to be taken on the current state and reward.
- From this point, the tasks that machine learning needs to complete are derived, referred to as:

Learning tasks.

What Type

- It is related to the “methods” taken by machine learning algorithms:
  - which machine learning model to use,
  - how to improve the ability of prediction,
  - how to improve the accuracy of analysis,
  - how to interact with the environment.
- From this point, the paradigms that machine learning needs to adopt are derived, referred to as:

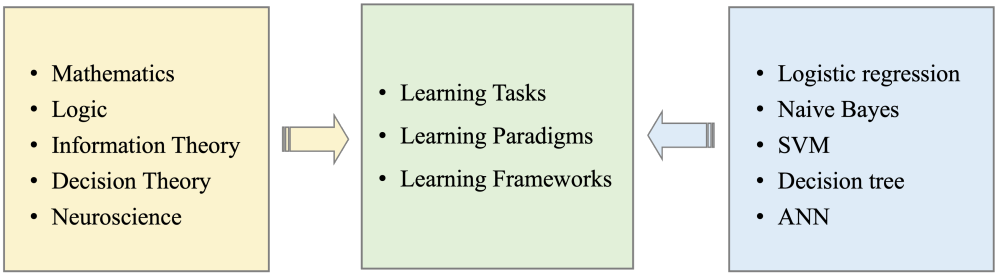
Learning paradigms.

How to Do

- It also needs to consider:
  - how to use suitable theories to design machine learning algorithms,
  - from which the theoretical framework that machine learning algorithms depend on.
- From this point, the frameworks that machine learning needs to based on are derived, referred to as:

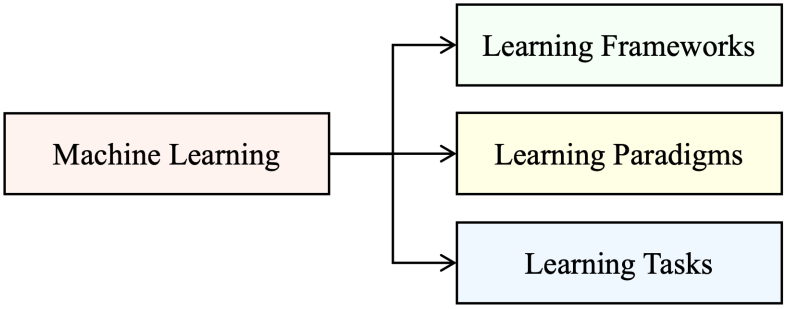
Learning frameworks.

Foundations, Algorithms, and Perspectives



The perspectives come from foundations and algorithms of machine learning.

The Three Perspectives



The Three Perspectives

Types	Description
Learning Frameworks	Denoting the different theories that can handle to fulfil some learning tasks.
Learning Paradigms	Denoting the different approaches that are happened in machine learning.
Learning Tasks	Denoting the different problems that can be solved by machine learning.

Learning Frameworks

What are Learning Frameworks

A learning framework refers to a theoretical framework for the design of machine learning algorithms. It belongs to the theory level of machine learning rather than the data level or task level, and it is also different from the software framework used to build machine learning algorithms.

Learning Frameworks

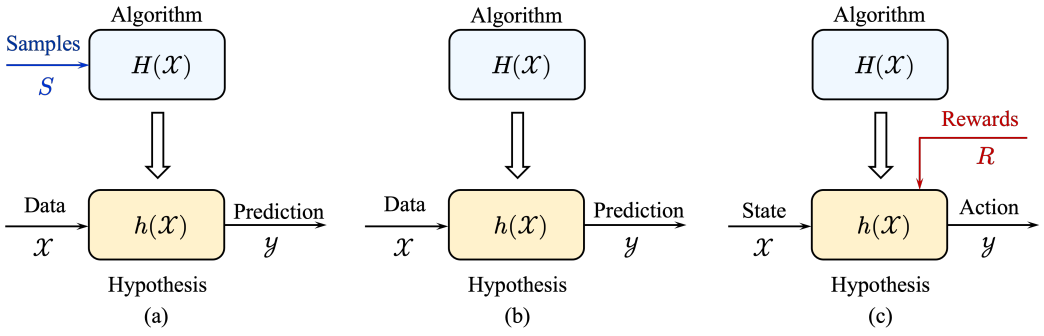
Framework	Brief Statements	Sub-frameworks
Probabilistic	Using probability to represent the conditional correlation of random variables.	Generative
		Discriminative
Statistical	Using statistics for constructing machine learning algorithms.	Parametric
		Non-parametric
Geometric	Using Euclidean or non-Euclidean geometry for data or algorithms.	Hyperplane
		Manifold
		Point Cloud
Connectionist	Based on the connectionist model of artificial neural networks.	Shallow/Deep
		Forward/Backward
Symbolic	Based on the physical symbolic system.	Logic-based
		Rule-based
		Causality-based
Behavioral	Based on the behavioral decision theory.	Multi-step
		Sequential

Learning Paradigms

What are Learning Paradigms

A learning paradigm in machine learning refers to the pattern or style taken by learning algorithms. It is a division of the paradigm of machine learning, independent of specific applications. The main basis to distinguish a learning paradigm is how to learn from data, or how to interact with the environment.

Learning Paradigms



Learning Paradigms

Paradigm	Brief Statements
Supervised	The algorithm H is trained with manually labeled samples S, and the obtained hypothesis h is then used to predict all unknown data.
Unsupervised	The algorithm H is trained only with unlabeled data, and the obtained h is then used to process all unknown data.
Reinforcement	The algorithm H is optimized to obtain h, and then the next action is determined based on the current state and the reward R of the environment.

Learning Tasks

What are Learning Tasks

A learning task refers to a basic problem that can be solved by machine learning. It is a common problem abstracted from domain problems, and the algorithms of machine learning with the same mechanism can be used to complete this task.

Learning Tasks

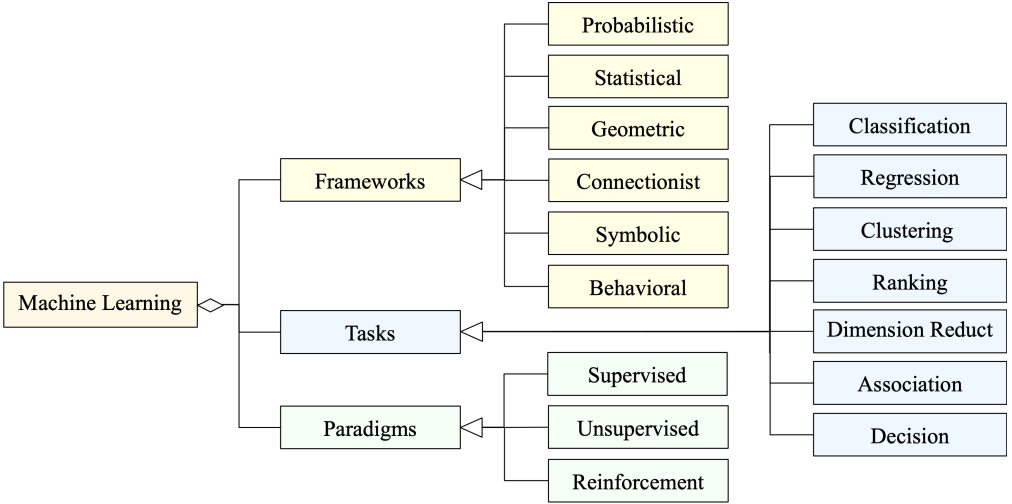
Paradigm	Brief Statements	Typical algorithm
Classification	Dividing input data into known categories.	SVM
Regression	Outputting continuous rather than discrete value.	Ridge regression
Clustering	Dividing input into groups unknown in advance.	k-Means
Ranking	Transforming order of input data to their rankings.	PageRank
Dimensionality reduction	Mapping high-dimensional data to low-dimensional.	PCA
Association	Discovering meaningful relationships between data.	Association rule
Decision	Including multistep and sequential decisions.	Q-Learning

Logical and Hierarchical Relationships

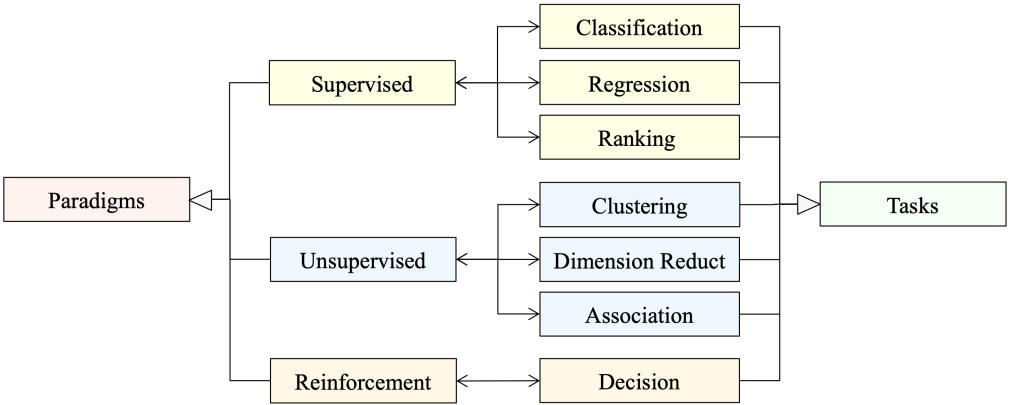
- Through in-depth analysis of the three perspectives and their related frameworks, paradigms, and tasks, we can find that they contain:
  - logical relationships,
  - hierarchical relationship.

- The logical relationships can be further divided into:
  - logical relationship of the three perspectives,
  - logical relationship between paradigms and tasks.

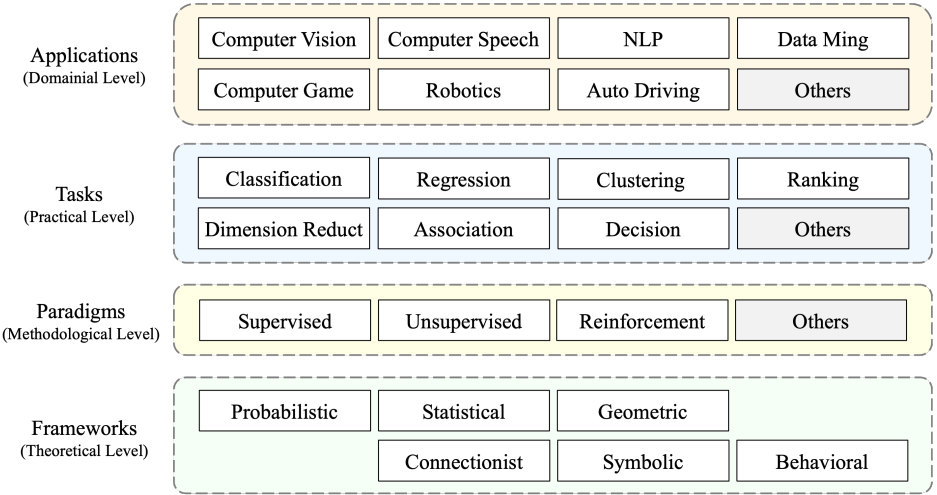
Logical Relationship of the Three Perspectives



Logical Relationship between Paradigms and Tasks



Hierarchical Relationship



Thank You