

Machine Learning

What is Learning?

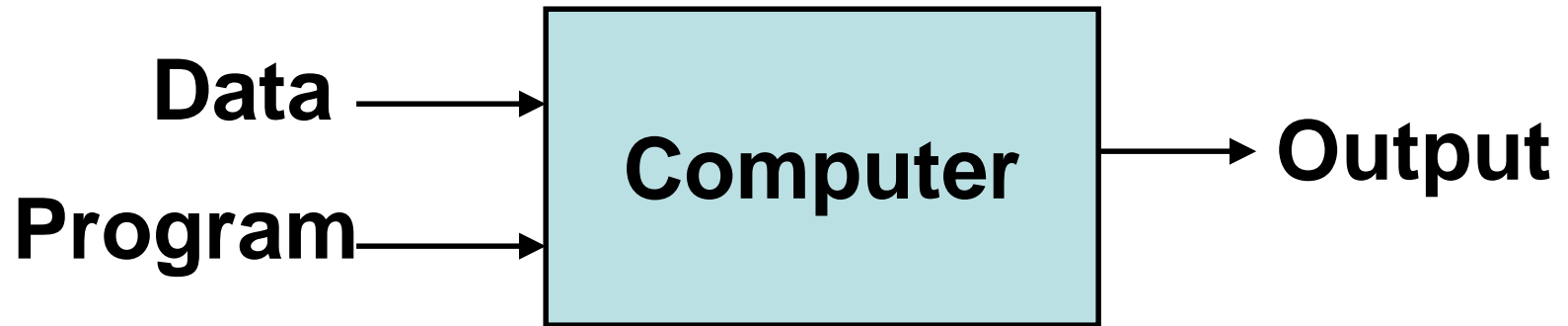
- Herbert Simon: “Learning is any process by which a system improves performance from experience.”
- “A computer program is said to **learn** from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E .”
– Tom Mitchell

What Is Machine Learning?

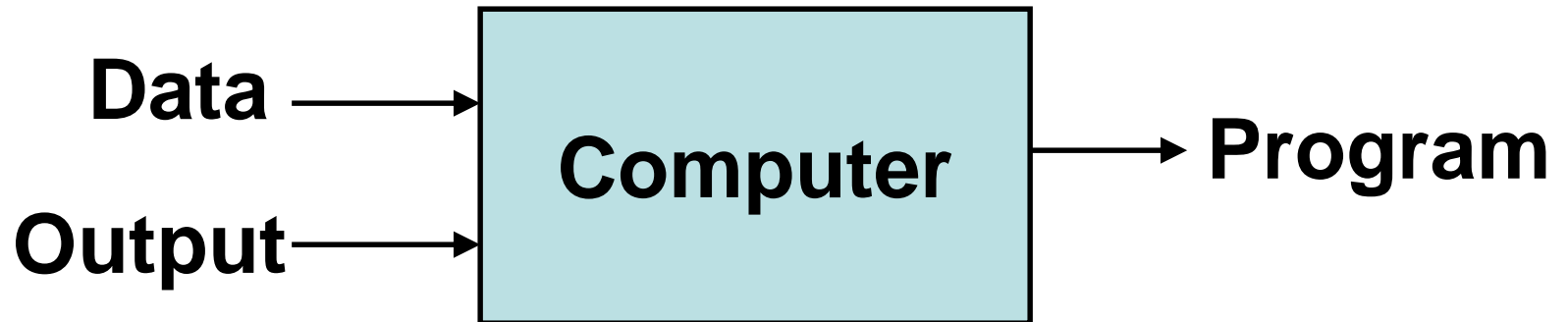
"Field of study that gives computers the ability to learn without being explicitly programmed".

-Arthur L. Samuel

Traditional Programming



Machine Learning



Machine Learning.....

- Machine learning is about designing algorithms that allow a computer to learn from experience and data.

Machine learning is used in cases where:

- There is an intuition that a certain rule exists
- But, we do not know it or cannot express it mathematically

So, we learn the rules from data

An Example



A bank wants to know whether to assign loan to a person or not

ID	Age	Has_Job	Own_House	Credit_Rating	Class
1	young	false	false	fair	No
2	young	false	false	good	No
3	young	true	false	good	Yes
4	young	true	true	fair	Yes
5	young	false	false	fair	No
6	middle	false	false	fair	No
7	middle	false	false	good	No
8	middle	true	true	good	Yes
9	middle	false	true	excellent	Yes
10	middle	false	true	excellent	Yes
11	old	false	true	excellent	Yes
12	old	false	true	good	Yes
13	old	true	false	good	Yes
14	old	true	false	excellent	Yes
15	old	false	false	fair	No

Bank has these data of previous years

An Example



A bank wants to know whether to assign loan to a person or not

Learn a classification model from the data

Use the model to classify future loan applications into

Yes (approved) and

No (not approved)

What is the class for following case/instance?

Age	Has_Job	Own_house	Credit-Rating	Class
young	false	false	good	?

An Example

Data: Loan application data

Task: Predict whether a loan should be approved or not.

Performance measure: accuracy.

No learning: classify all future applications (test data) to the majority class (i.e., **Yes**):

Accuracy = 9/15 = 60%.

We can do better than 60% with learning.

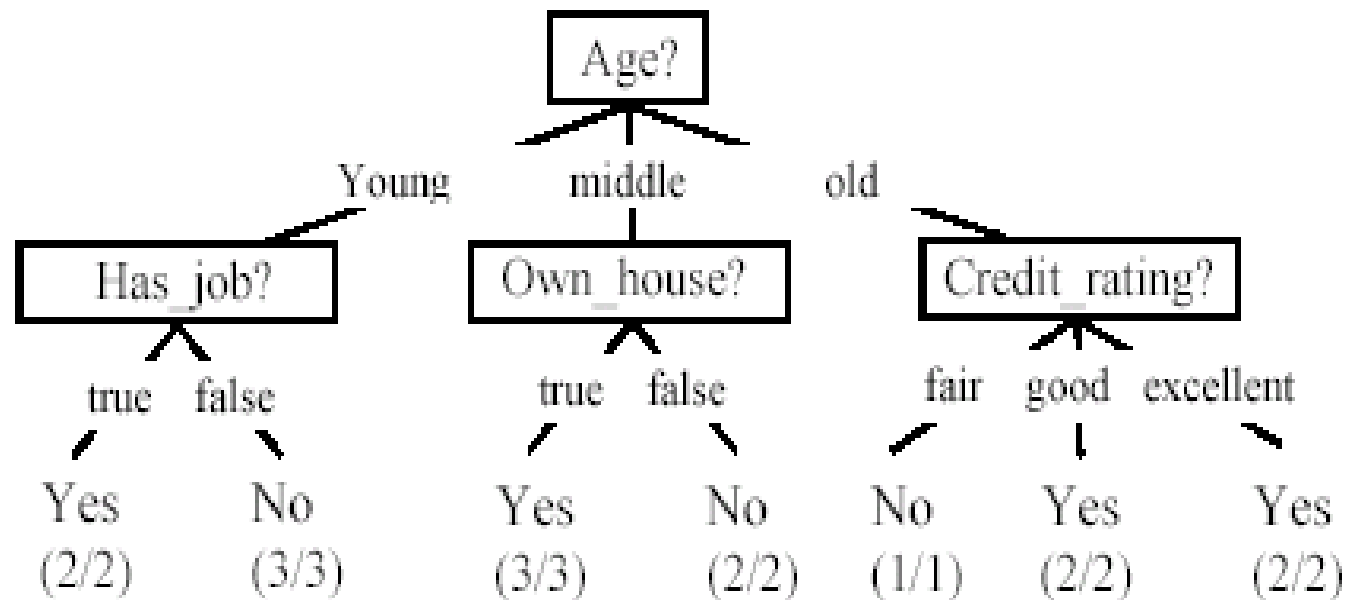
ID	Age	Has_Job	Own_House	Credit_Rating	Class
1	young	false	false	fair	No
2	young	false	false	good	No
3	young	true	false	good	Yes
4	young	true	true	fair	Yes
5	young	false	false	fair	No
6	middle	false	false	fair	No
7	middle	false	false	good	No
8	middle	true	true	good	Yes
9	middle	false	true	excellent	Yes
10	middle	false	true	excellent	Yes
11	old	false	true	excellent	Yes
12	old	false	true	good	Yes
13	old	true	false	good	Yes
14	old	true	false	excellent	Yes
15	old	false	false	fair	No

Age	Has_Job	Own_house	Credit-Rating	Class
young	false	false	good	?

An Example

A decision tree approach

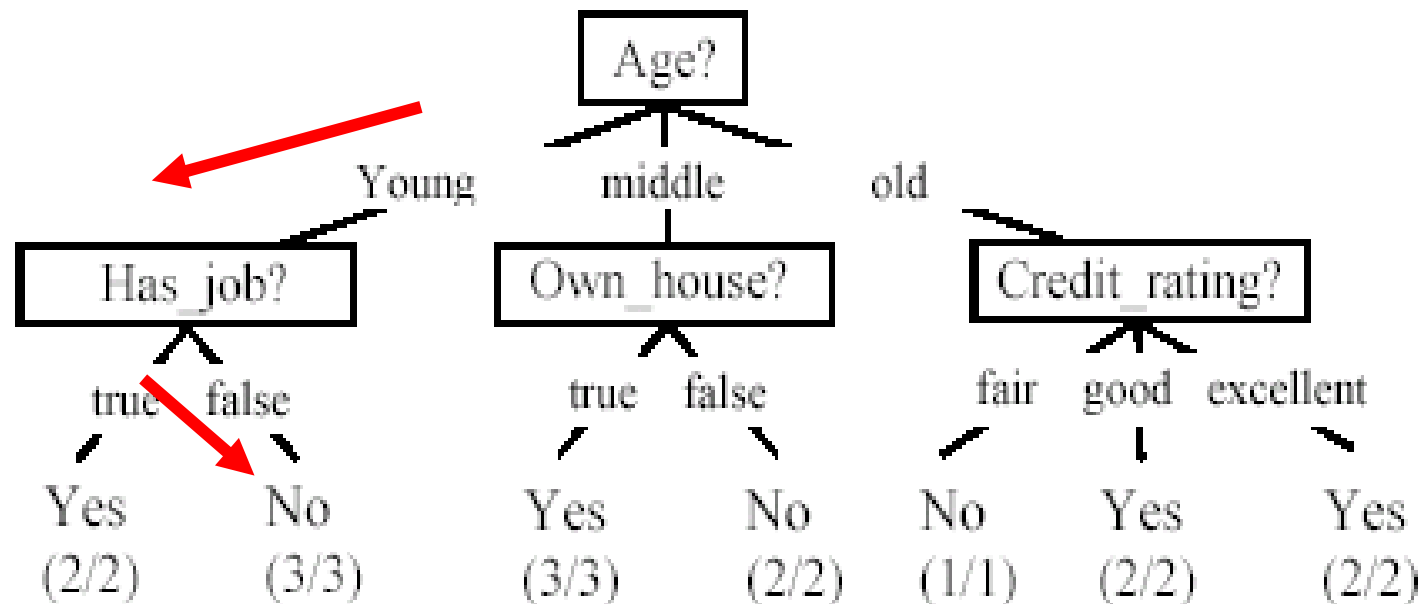
ID	Age	Has_Job	Own_House	Credit_Rating	Class
1	young	false	false	fair	No
2	young	false	false	good	No
3	young	true	false	good	Yes
4	young	true	true	fair	Yes
5	young	false	false	fair	No
6	middle	false	false	fair	No
7	middle	false	false	good	No
8	middle	true	true	good	Yes
9	middle	false	true	excellent	Yes
10	middle	false	true	excellent	Yes
11	old	false	true	excellent	Yes
12	old	false	true	good	Yes
13	old	true	false	good	Yes
14	old	true	false	excellent	Yes
15	old	false	false	fair	No



An Example

A decision tree approach

Age	Has_Job	Own_house	Credit-Rating	Class
young	false	false	good	No



Sample Applications

- Web search
- Computational biology
- Finance
- E-commerce
- Space exploration
- Robotics
- Information extraction
- Social networks
- ...

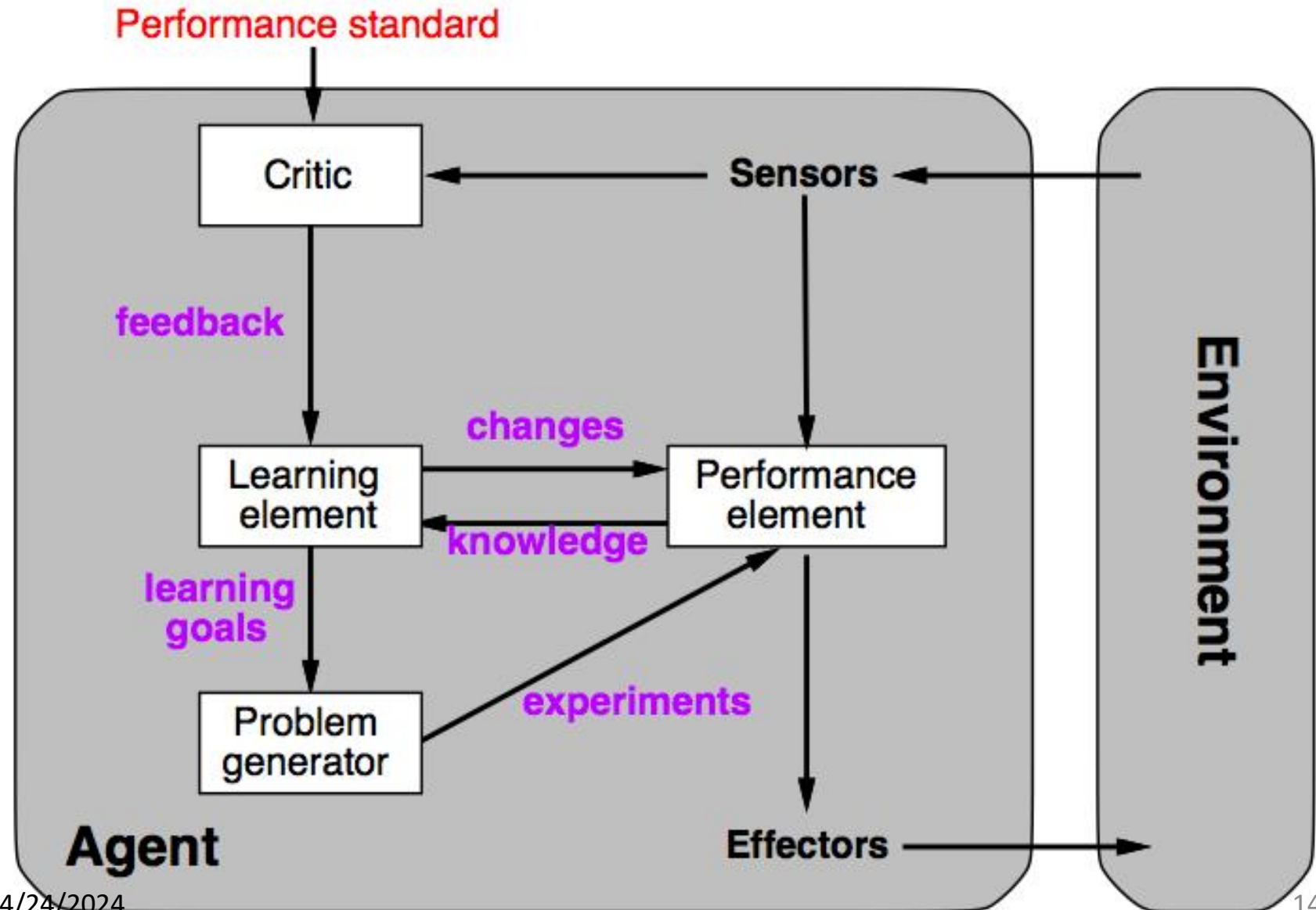
Representation

- Decision trees
- Sets of rules / Logic programs
- Instances
- Graphical models (Bayes/Markov nets)
- Neural networks
- Support vector machines
- ...

Evaluation

- Accuracy
- Precision and recall
- Squared error
-
-
-

Structure of a learning agent



Machine Learning Areas

- **Supervised Learning:** Data and corresponding labels are given
- **Unsupervised Learning:** Only data is given, no labels provided
- **Semi-supervised Learning:** Some (if not all) labels are present
- **Reinforcement Learning:** An agent interacting with the world makes observations, takes actions, and is rewarded or punished; it should learn to choose actions in such a way as to obtain a lot of reward

Classification—A Two-Step Process

- **Model construction**: describing a set of predetermined classes
 - Each tuple/sample is assumed to belong to a predefined class, as determined by the **class label**
 - The set of tuples used for model construction is **training set**
 - The model is represented as **classification rules, decision trees, or mathematical formulae**
- **Model usage**: for classifying future or unknown objects
 - **Estimate accuracy** of the model
 - The known label of test sample is compared with the classified result from the model
 - **Test set is independent of training set**, otherwise overfitting will occur
 - If the accuracy is acceptable, use the model to **classify data** tuples whose class labels are not known

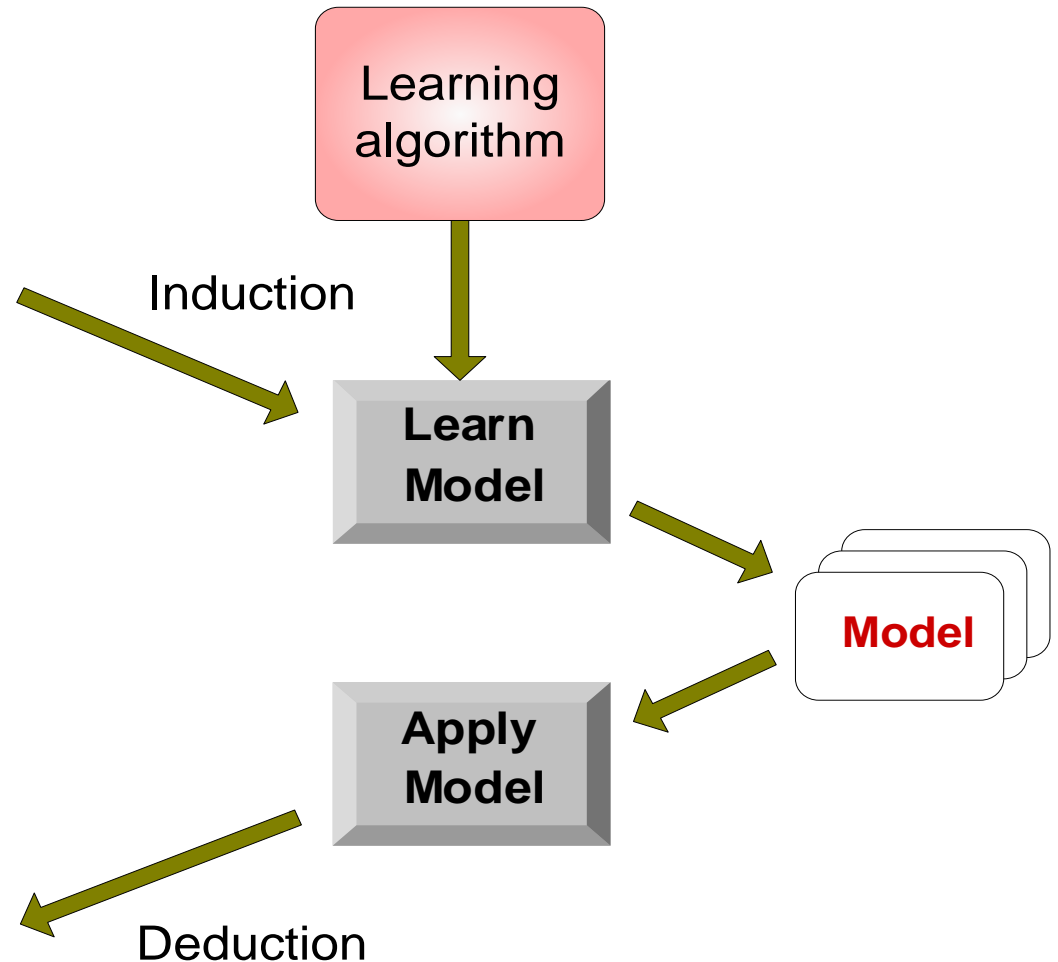
Illustrating Classification Task

Tid	Attrib1	Attrib2	Attrib3	Class
1	Yes	Large	125K	No
2	No	Medium	100K	No
3	No	Small	70K	No
4	Yes	Medium	120K	No
5	No	Large	95K	Yes
6	No	Medium	60K	No
7	Yes	Large	220K	No
8	No	Small	85K	Yes
9	No	Medium	75K	No
10	No	Small	90K	Yes

Training Set

Tid	Attrib1	Attrib2	Attrib3	Class
11	No	Small	55K	?
12	Yes	Medium	80K	?
13	Yes	Large	110K	?
14	No	Small	95K	?
15	No	Large	67K	?

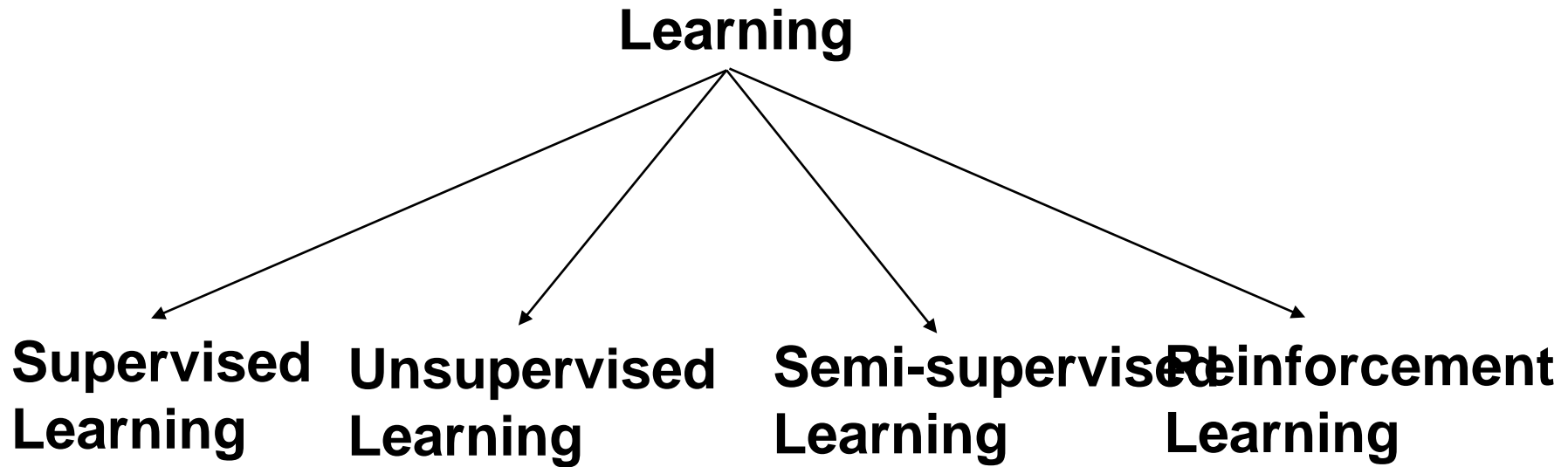
Test Set



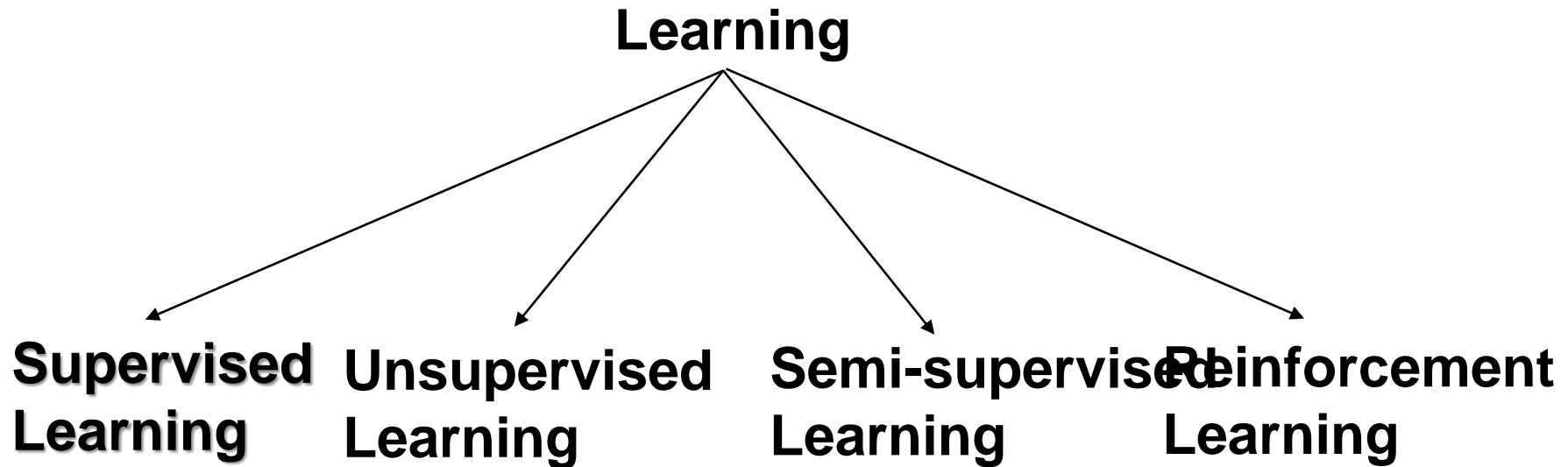
Classification Examples

- In classification, we predict labels y (classes) for inputs x
- Examples:
 - OCR (input: images, classes: characters)
 - Medical diagnosis (input: symptoms, classes: diseases)
 - Automatic essay grader (input: document, classes: grades)
 - Fraud detection (input: account activity, classes: fraud / no fraud)
 - Customer service email routing
 - Recommended articles in a newspaper, recommended books
 - DNA and protein sequence identification
 - Categorization and identification of astronomical images
 - Financial investments
 - ... many more

Types of learning



Types of learning



Supervised Learning

- There is a learner and a supervisor(teacher).
- Under the complete supervision of the teacher, the learner acquires the knowledge about the data.



Class: Dog

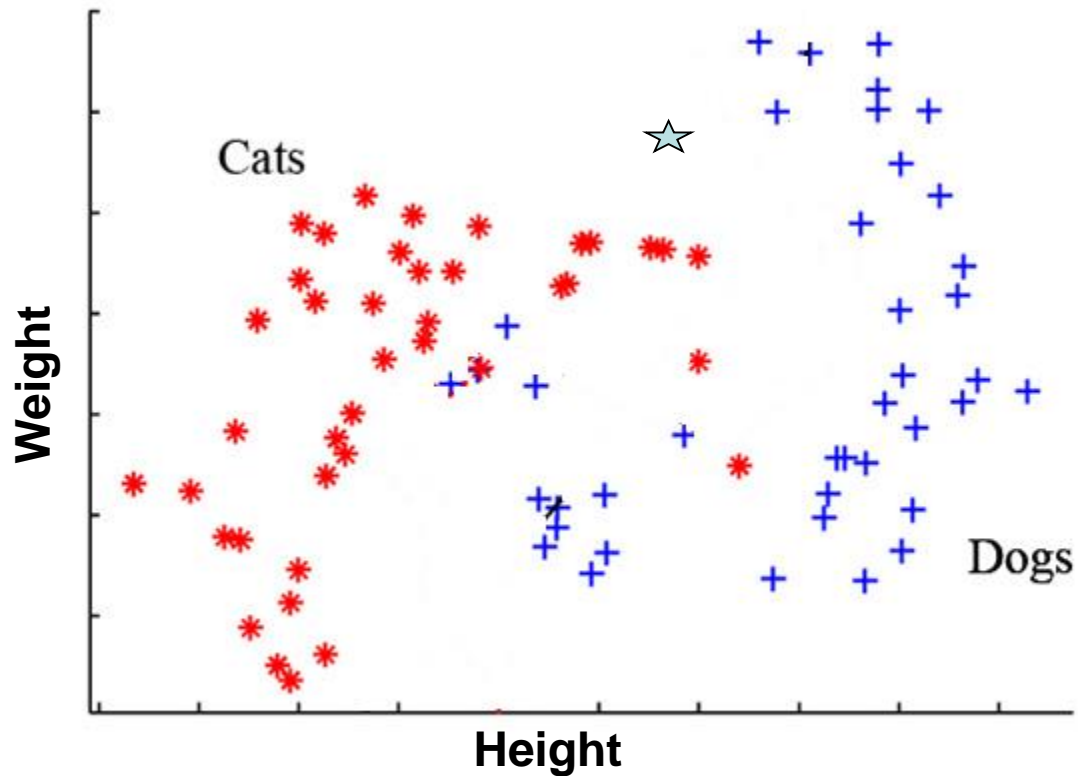


Class: Cat



Unknown test samples

Supervised Learning

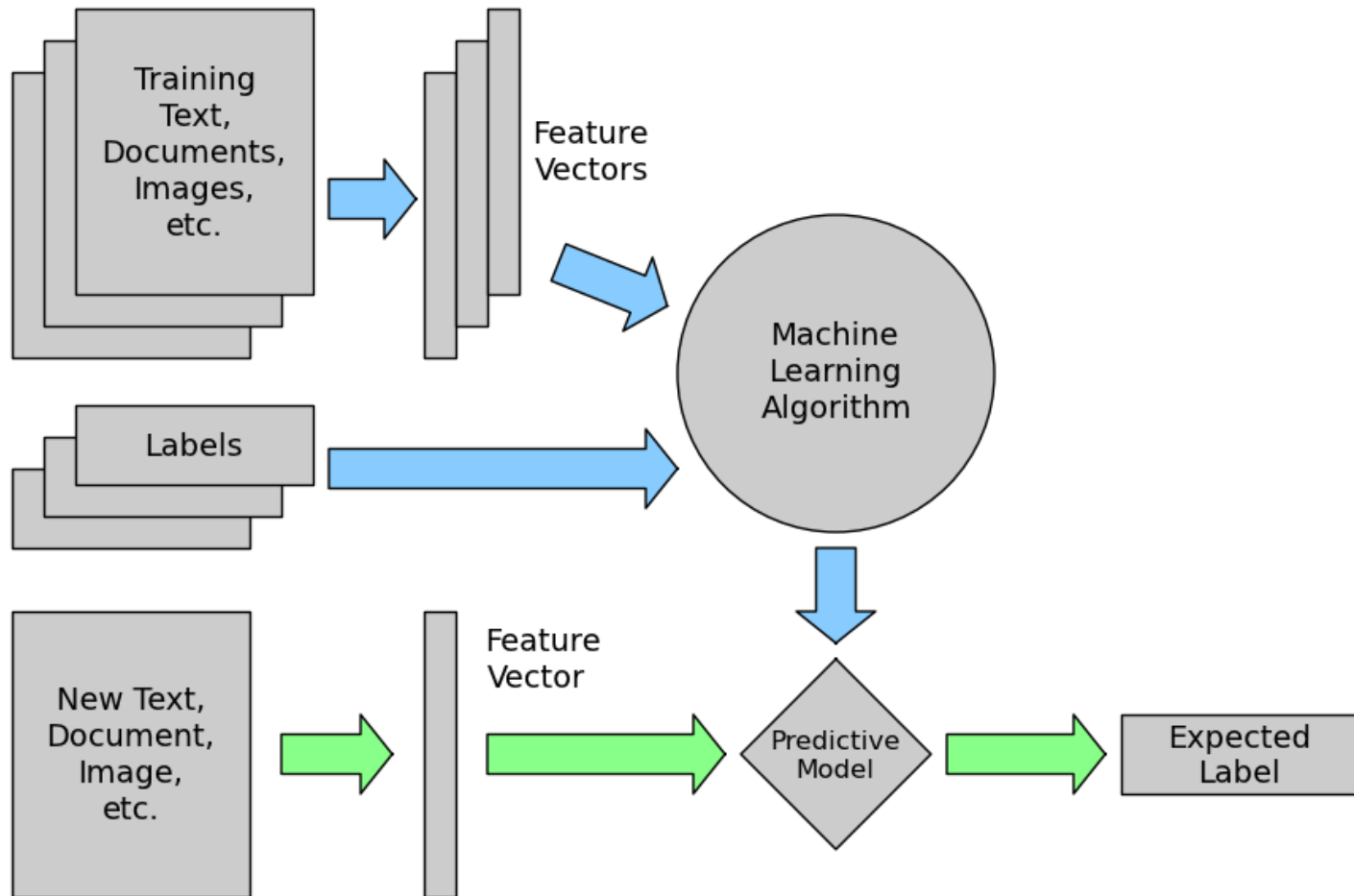


Classify the unknown test sample on the basis of the available training data

Supervised Learning

- The data is a collection of input-output patterns
- The objects under consideration are associated with target values or labels (which take the role of the teacher).
- These **labeled** examples constitute the **training** set.
- The algorithm generates a model that maps the inputs to their desired outputs.
- This model is used to predict output for unknown input samples.

Supervised Learning Model



Types of Supervised Learning

Classification

- **The PR problem of assigning an object to a class**
- **The output of the PR system is an integer label**

e.g. classifying a product as “good” or “bad” in a quality control test

Regression

- **A generalization of a classification task**
- **The output of the PR system is a real-valued number**

e.g. predicting the share value of a firm based on past performance and stock market indicators

Supervised Learning 1: Classification

- **Classification** is the problem of identifying the category to which a new observation belongs, on the basis of a training set of data containing observations whose category is known.
- The output space is discrete.
 - The inputs can correspond to only a set of output categories.

Classification

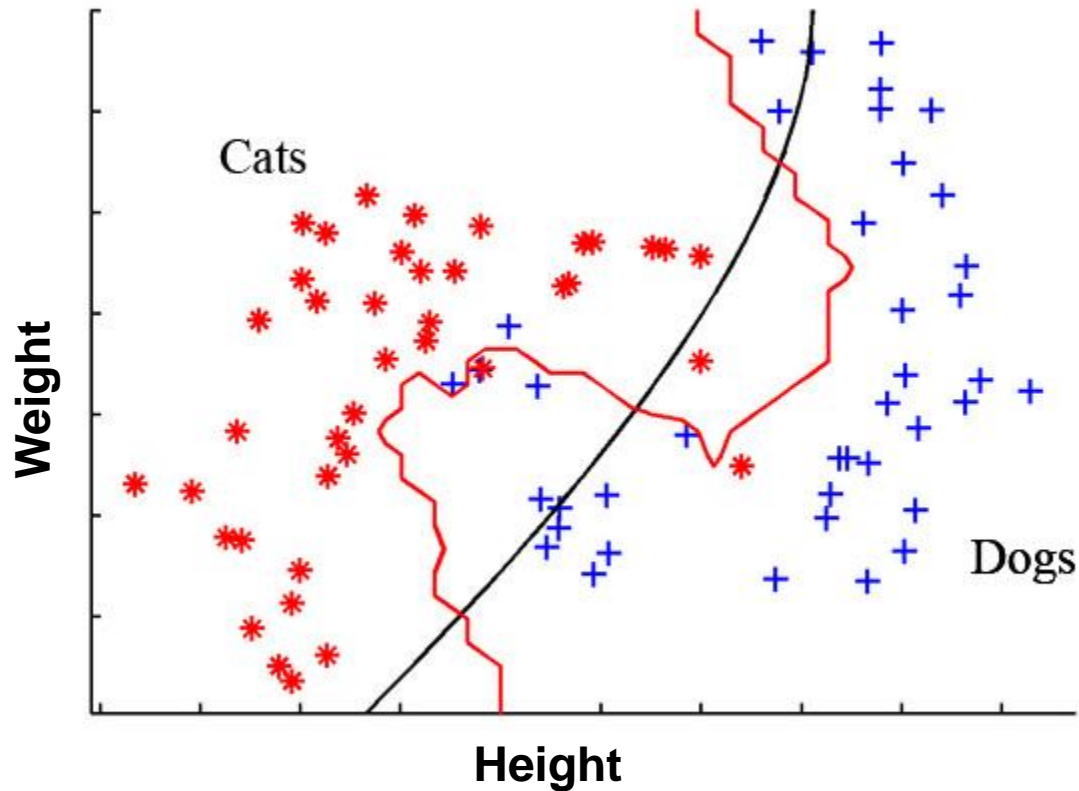


Cats

Dogs



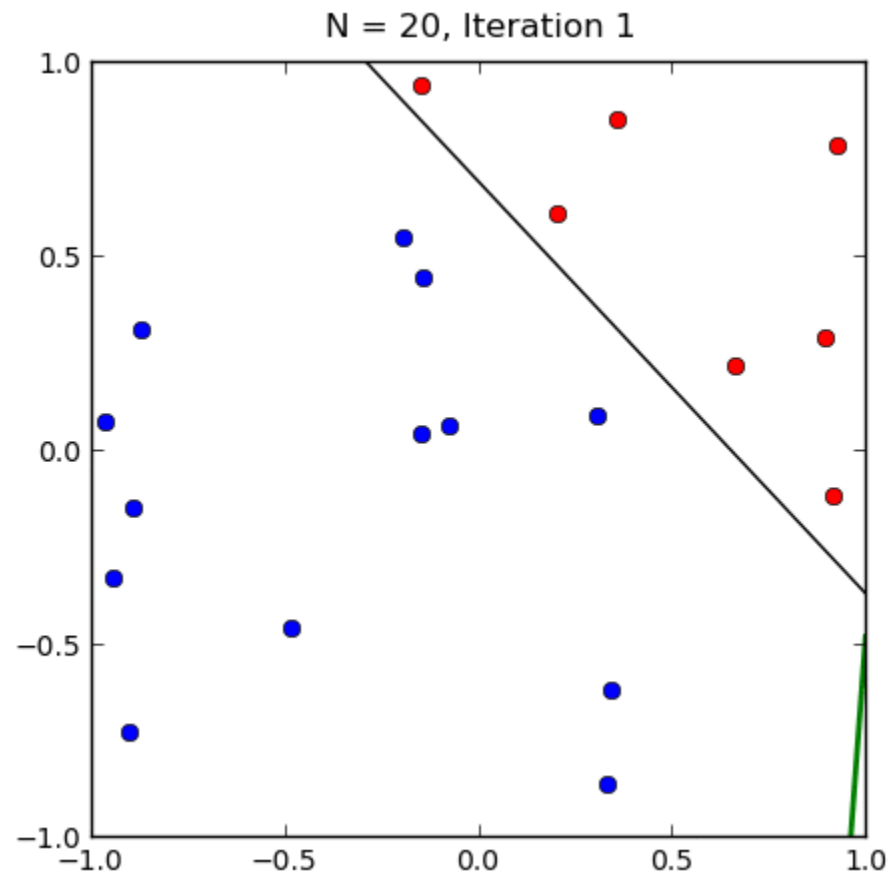
Classification



A decision boundary is chosen to separate the classes in a best possible way

Classification

Here, a classification algorithm tries to find the best possible decision boundary that can separate the two classes.



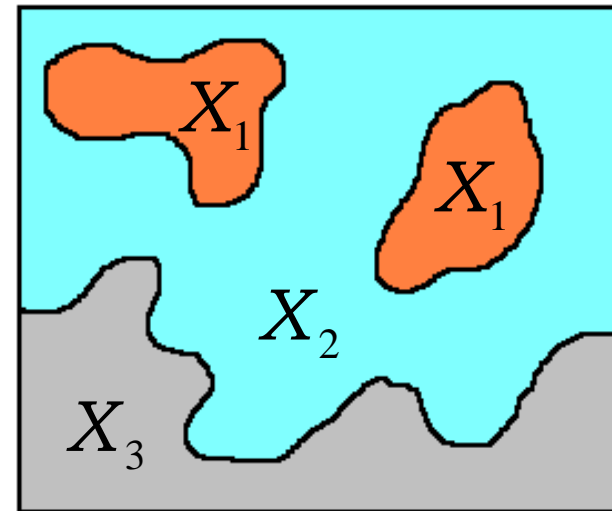
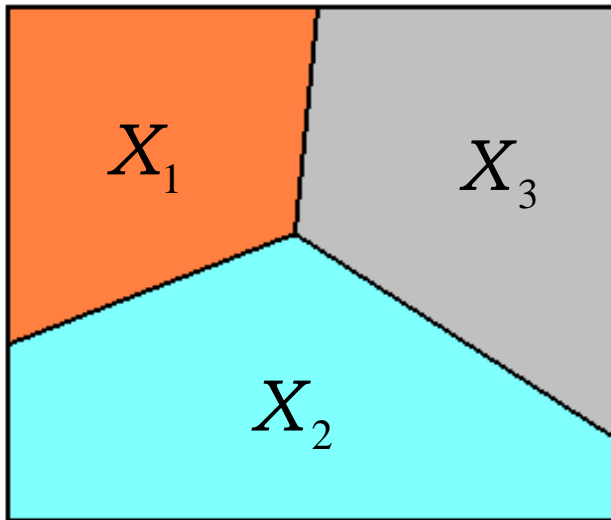
Classification

- When there are only two output choices, this is called **two-class** or **binomial** classification.
- When there are more categories the problem is known as **multi-class** classification.

Clustering

A clustering partitions feature space X into non-overlapping regions such that

$$X = X_1 \cup X_2 \cup \dots \cup X_{|Y|} \quad \text{and} \quad X_1 \cap X_2 \cap \dots \cap X_{|Y|} = \{0\}$$

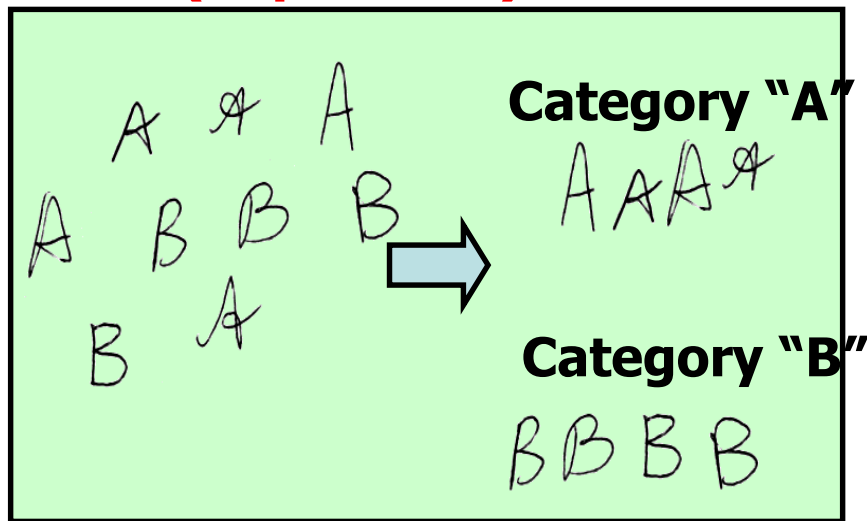


- Patterns in a group are similar than patterns between clusters.

Classification **vs** Clustering

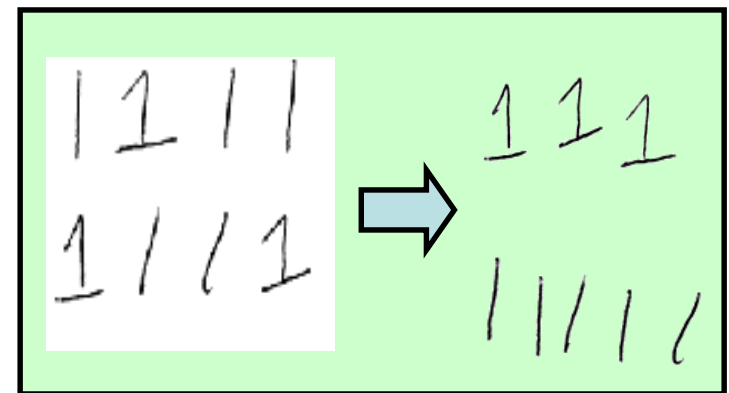
Classification (known categories)

(Supervised)



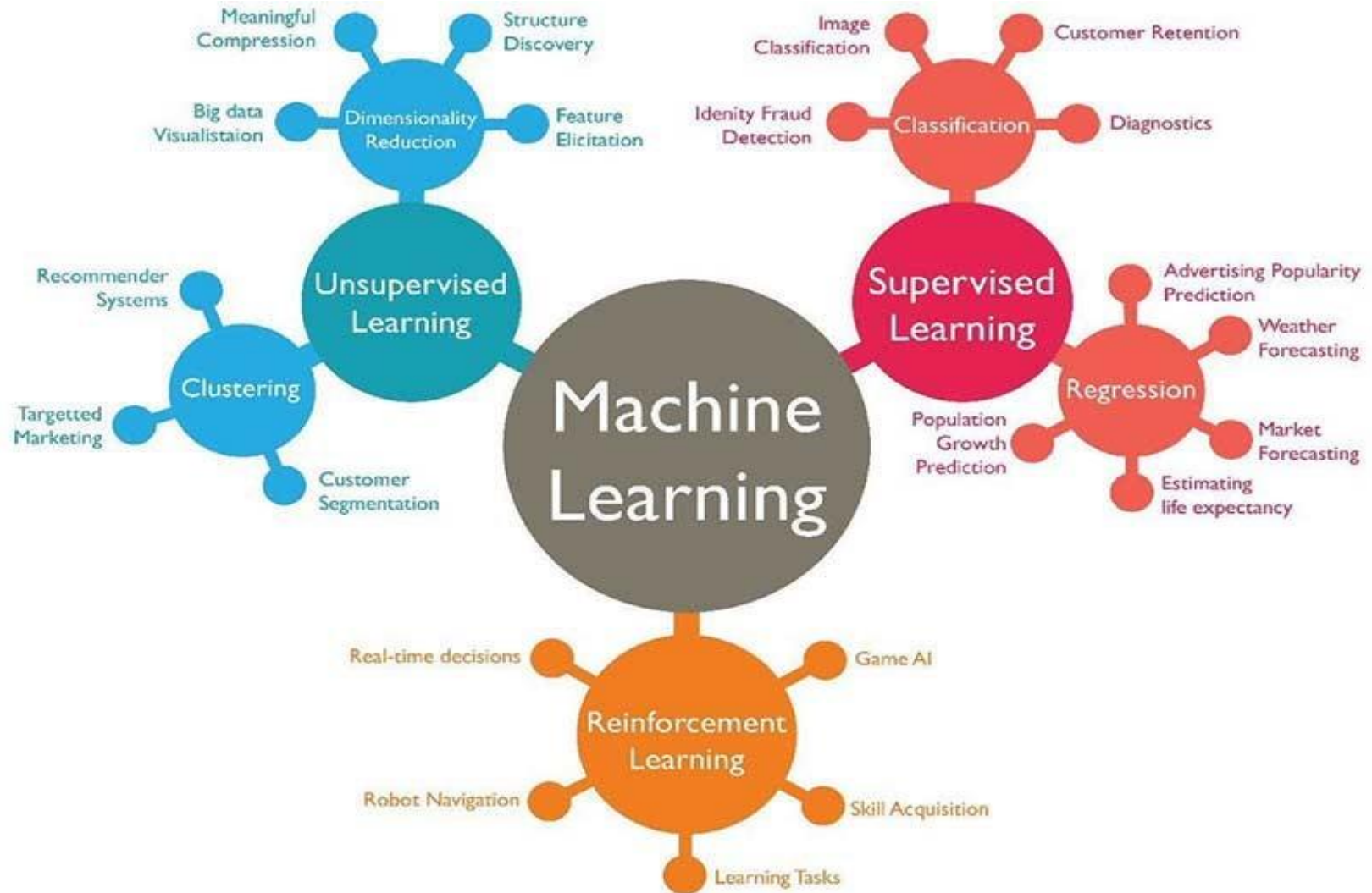
Clustering (unknown categories)

(Unsupervised)



Good Clustering

- **A good clustering method will produce high quality clusters with**
 - **high intra-class similarity**
 - **low inter-class similarity**
- **The quality of a clustering result depends on both the similarity measure used by the method and its implementation.**
- **The quality of a clustering method is also measured by its ability to discover some or all of the hidden patterns.**



References

Books

Duda, Hart: Pattern Classification and Scene Analysis. J. Wiley & Sons, New York, 1982. (2nd edition 2000).

Bishop: Neural Networks for Pattern Recognition. Claredon Press, Oxford, 1997.