Machine Learning

Lecture 2-3: Types of Machine Learning

COURSE CODE: CSE451

2023

Course Teacher
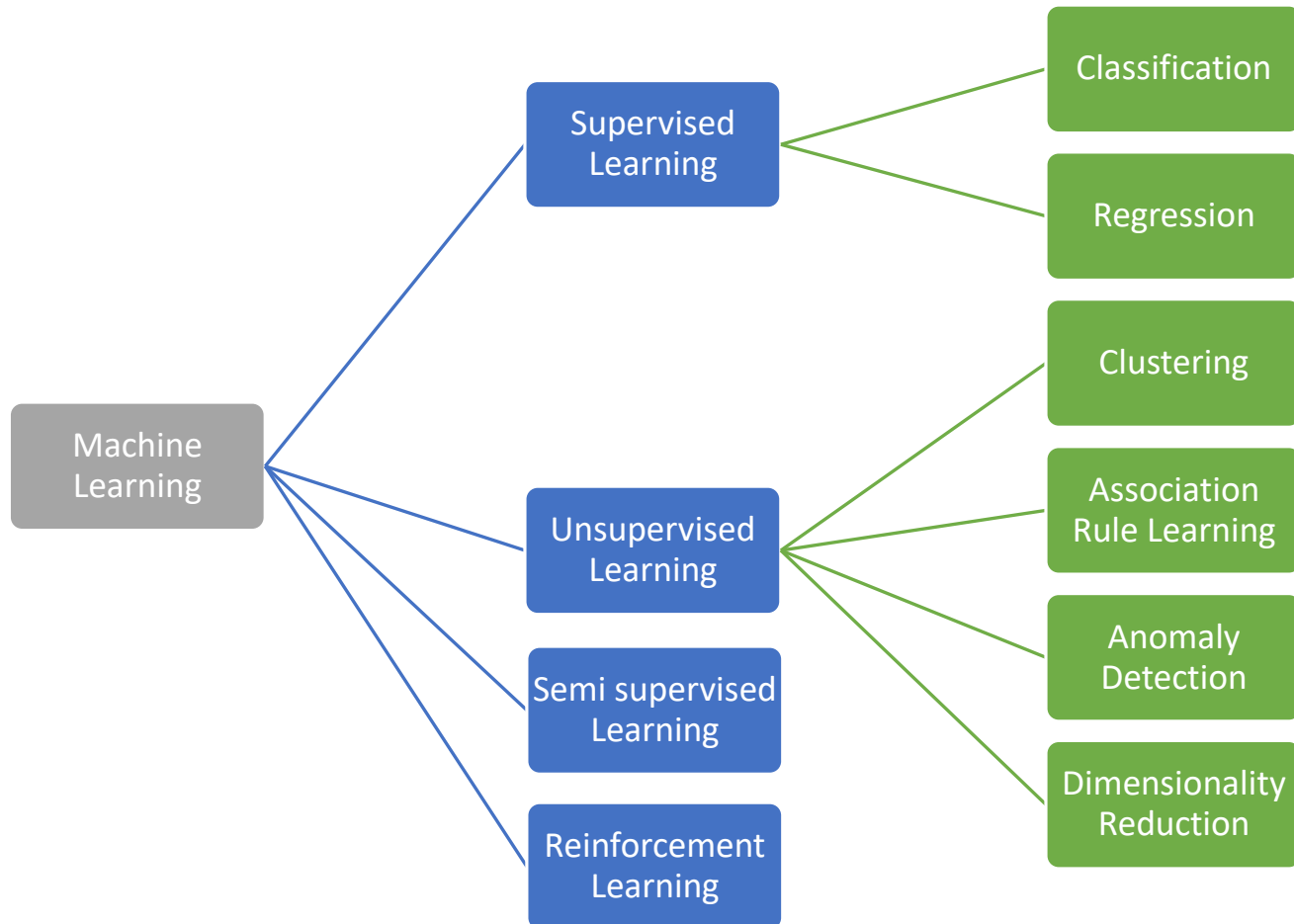
Dr. Mrinal Kanti Baowaly

Associate Professor

Department of Computer Science and Engineering, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Bangladesh.

Email: mkbaowaly@gmail.com

Types of Machine Learning

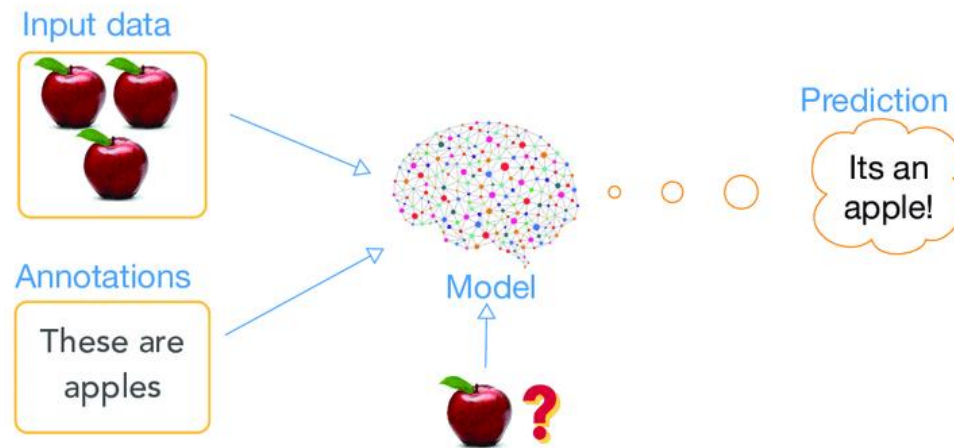# Supervised Learning vs Unsupervised Learning (An analogy)
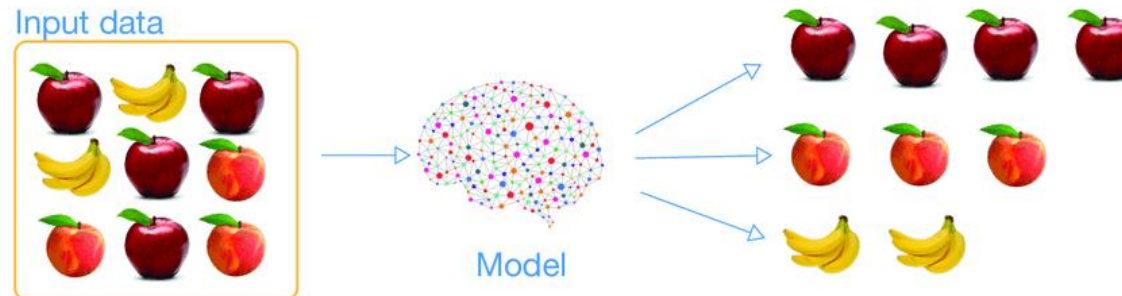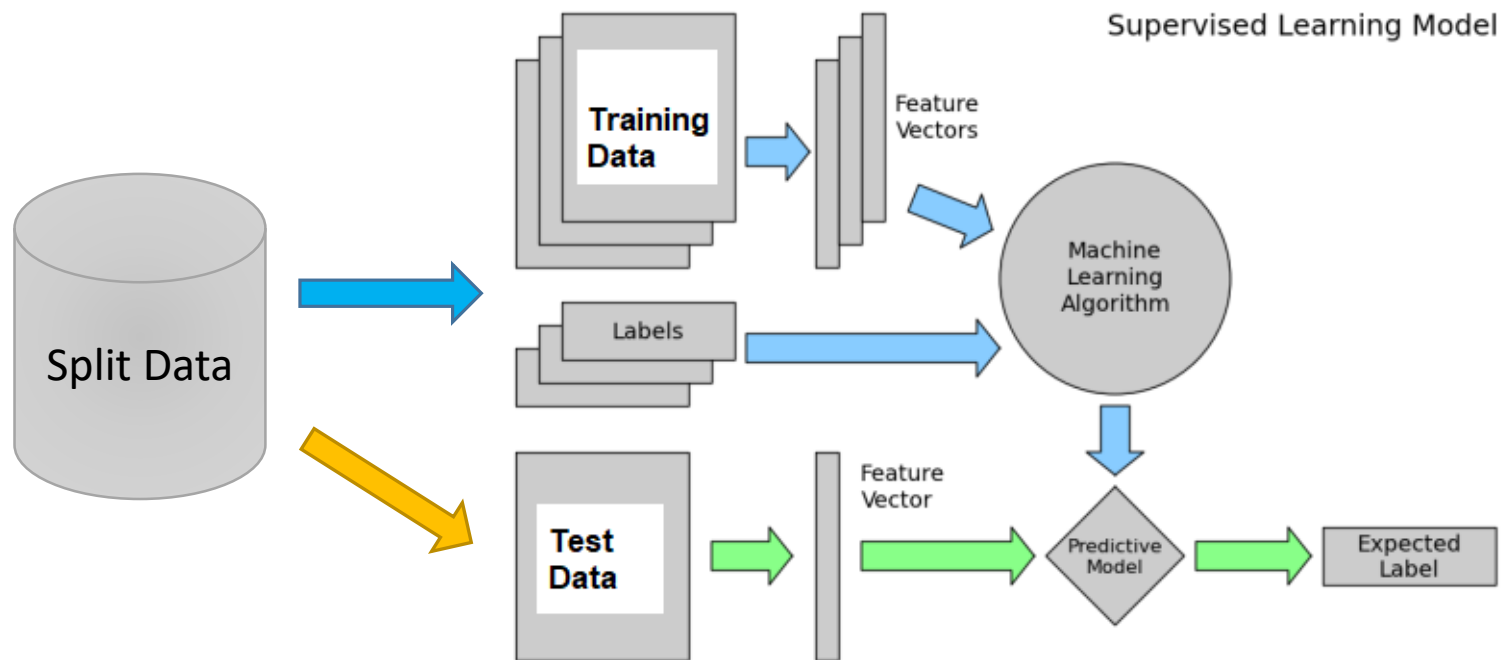
Supervised Learning

- Also called Predictive Modeling

Supervised Learning (Cont.)

- Supervised learning is used to learn a mapping function f from the input *(X)* to the output *Y* as follows:

  Y = f(X)

- The goal is to approximate the mapping function f so well that when you have new input data *(X)* that it can predict the output variables *(Y)* for that data.

- It is called supervised learning because the process of an algorithm learning from the training dataset can be thought of as a teacher supervising the learning process

Classification

- When the output variable / target variable is a category or a set of categories, such as "disease or no disease", "red or green or blue" etc.

- Predicts one or more categories/labels/classes for each input/instance

- Binary Classification: Classifying instances into one of two classes/categories, for example, email spam filtering (spam or not).

- Multiclass Classification: Classifying instances into one of three or more classes/categories, for example, color of an object, it may be red or green or blue.

- Multi-Label Classification: Multiple labels/classes are to be predicted for each instance, for example, when predicting a given movie category, it may belong to horror, romance, adventure, action, or all simultaneously.

Some Classification Problems

- Spam filtering

- Language detection

- Document/text classification

- Sentiment analysis of text (positive, negative, or neutral)

- Recognition of handwritten characters and numbers

- Fraud detection

- Disease prediction and detection etc.

Classification
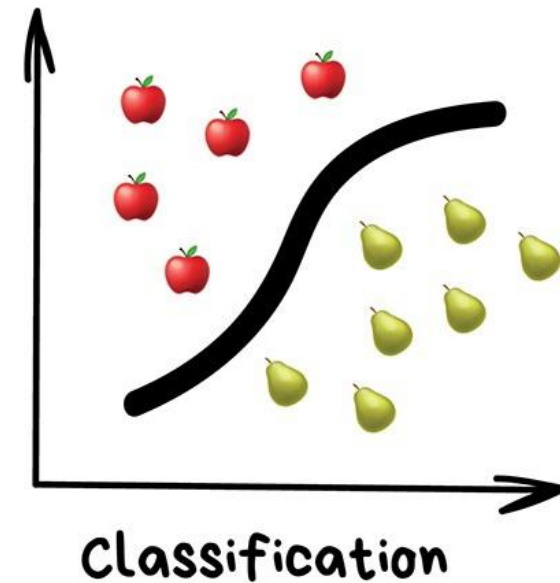
Regression

- When the output variable is a real/continuous value, such as "dollars" or "weight" or "Score"

- Predicts a single output value

- Why do we use Regression Analysis?
  - Forecasting
  - Demand and sales volume analysis
  - Time series modelling
  - Medical diagnosis etc.

- Linear Regression

- Polynomial Regression
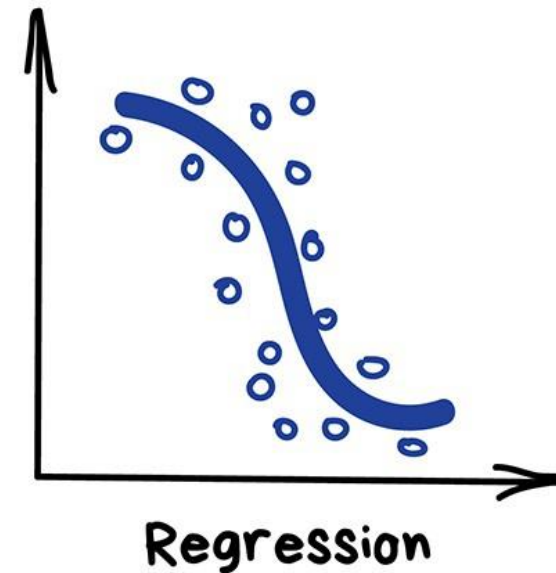
Source: AnalyticsVidhya

List of Commonly Supervised Learning Algorithms

- Linear Regression
- Logistic Regression
- k-Nearest Neighbors (kNN)
- Decision Trees
- Random Forest
- Support Vector Machines (SVM)
- Gradient Boosting Machines (GBM)
- LightGBM
- XGBoost
- CatBoost
- Neural Networks

What is Unsupervised Learning?

- Unsupervised learning is where you only have unlabeled input data (X) and allow the algorithm to work on its own to discover the underlying groupings, structure or pattern in the data.

- These are called unsupervised learning because unlike supervised learning there is no correct answers and there is no teacher (i.e., learning from the labeled training data).

- Unsupervised learning problems can be further grouped into **clustering, association rule mining**, **anomaly detection**, and **dimensionality reduction**.

What is Cluster Analysis / Clustering?

Finding groups of objects such that the objects in a group will be similar (or related) to one another and different from (or unrelated to) the objects in other groups
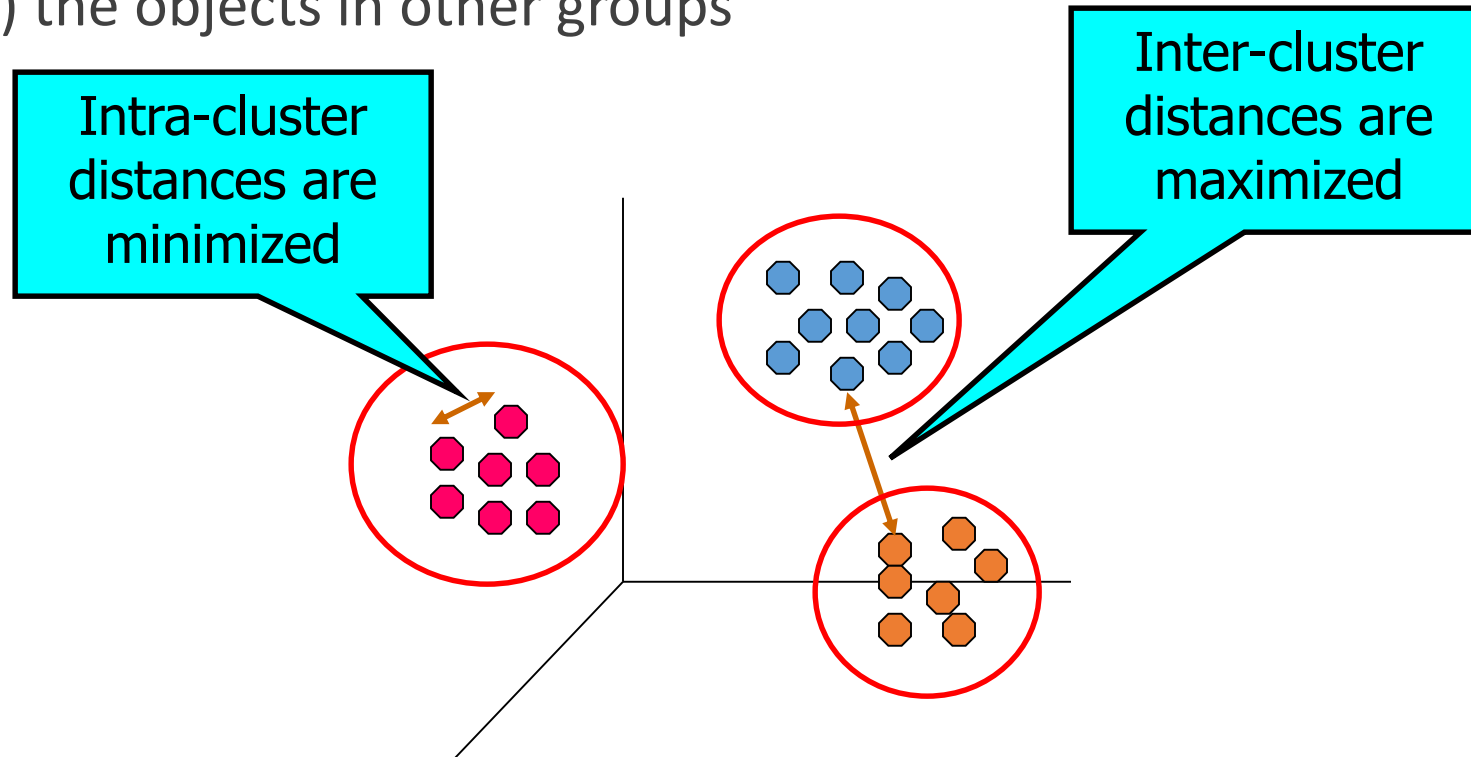
Applications of Clustering

- **Customer Segmentation:** This strategy is across functions, including banking, telecom, e-commerce, sports, advertising, sales, etc.

- **Document Clustering:** Cluster similar documents together

- **Image Clustering:** You can group similar images together.

- **Image Segmentation:** You can apply clustering to create clusters having similar pixels in the image together.

- **Recommendation Engines:** You can look at the songs liked by a person and then use clustering to find similar songs and finally recommend the most similar songs to him.

What is Association Rule Mining?

- Association Rule Mining is a rule-based machine learning method for discovering frequently occurring **patterns**, **correlations**, or **associations** between variables in large databases.

- It is intended to identify strong rules discovered in databases.

- A typical example is Market Based Analysis: for example, if a customer buys bread, he most likely can also buy butter, eggs, or milk, so these products are stored within a shelf or mostly nearby

Applications of Association Rule Mining

- **Market Basket Analysis:** It is one of the popular examples and applications of association rule mining. This technique is commonly used by big retailers to determine the association between items.

- **Medical Diagnosis:** With the help of association rules, patients can be cured easily, as it helps in identifying the probability of illness for a particular disease.

- **Protein Sequence:** The association rules help in determining the synthesis of artificial Proteins.

- **Census Data:** The association rule mining has immense potential in census data in supporting sound public policy and bringing forth an efficient functioning of a democratic society.

- It is also used for the **Catalog Design** and **Loss-leader Analysis** and many more other applications.   Source: javaTpoint, upGrad

List of Commonly Used Unsupervised Learning Algorithms

- K-means, Hierarchical for clustering problems

- Apriori algorithm for association rule learning problems

- Principal Component Analysis

- Singular Value Decomposition

- Independent Component Analysis

Supervised vs Unsupervised Learning

| Supervised Learning	Unsupervised Learning
Algorithms are trained using labeled data.	Algorithms are trained using unlabeled data.
The goal is to train the model so that it can predict the output when it is given new data.	The goal is to find underlying groupings, structure or pattern in the data.
Supervised learning model takes direct feedback to check if it is predicting correct output or not.	Unsupervised learning model does not take any feedback.
Supervised learning model produces an accurate result.	Unsupervised learning model may give less accurate result as compared to supervised learning.
Supervised learning is a simpler method.	Unsupervised learning is computationally complex.

Semisupervised learning

- Semisupervised learning can deal with partially labeled training data, usually a lot of unlabeled data and a little bit of labeled data.

- Most semisupervised learning algorithms are combinations of unsupervised and supervised algorithms.

- **Example**: Some photo-hosting services, such as Google Photos. Once you upload all your family photos to the service, it automatically recognizes that the same person A shows up in photos 1, 5, and 11, while another person B shows up in photos 2, 5, and 7 (unsupervised learning). Now all the system needs is for you to tell it who these people are (Supervised learning). Just one label per person, and it is able to name everyone in every photo, which is useful for searching photos.

Reinforcement Learning

- Reinforcement learning is a feedback-based learning method, in which a learning agent gets a reward for each right action and gets a penalty for each wrong action.

- The agent learns automatically with these feedbacks and improves its performance.

- The agent interacts with the environment and explores it.

- The goal of an agent is to get the most reward points, and hence, it improves its performance.

- Surviving in an environment is a core idea of reinforcement learning. For example, throw a robot into a maze and let it find an exit.
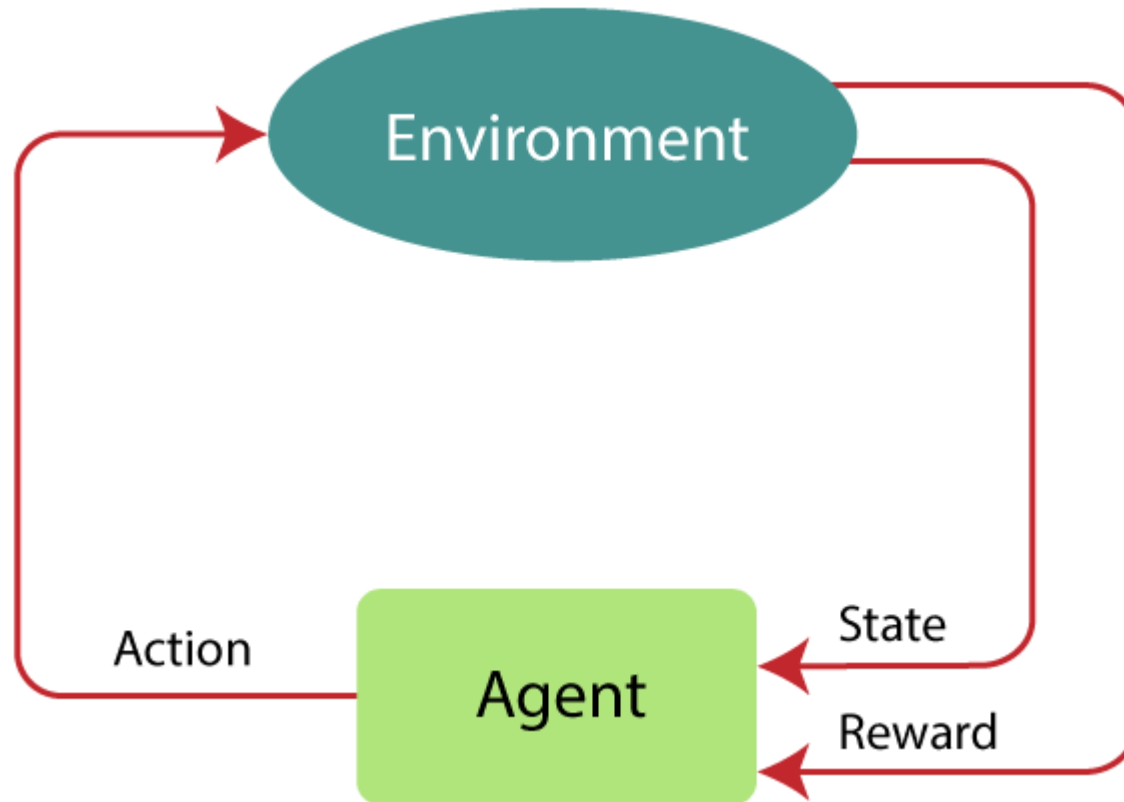
Reinforcement Learning (Analogies)
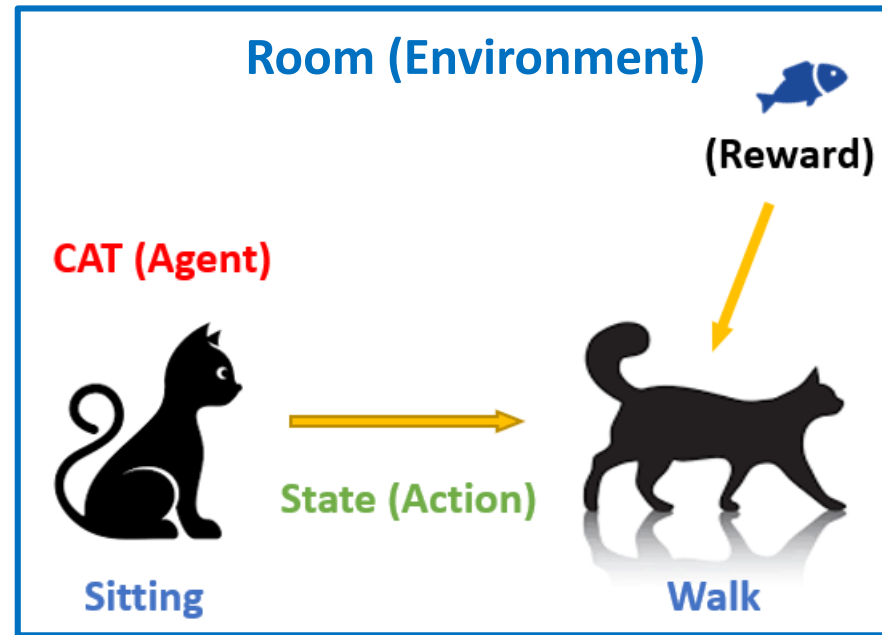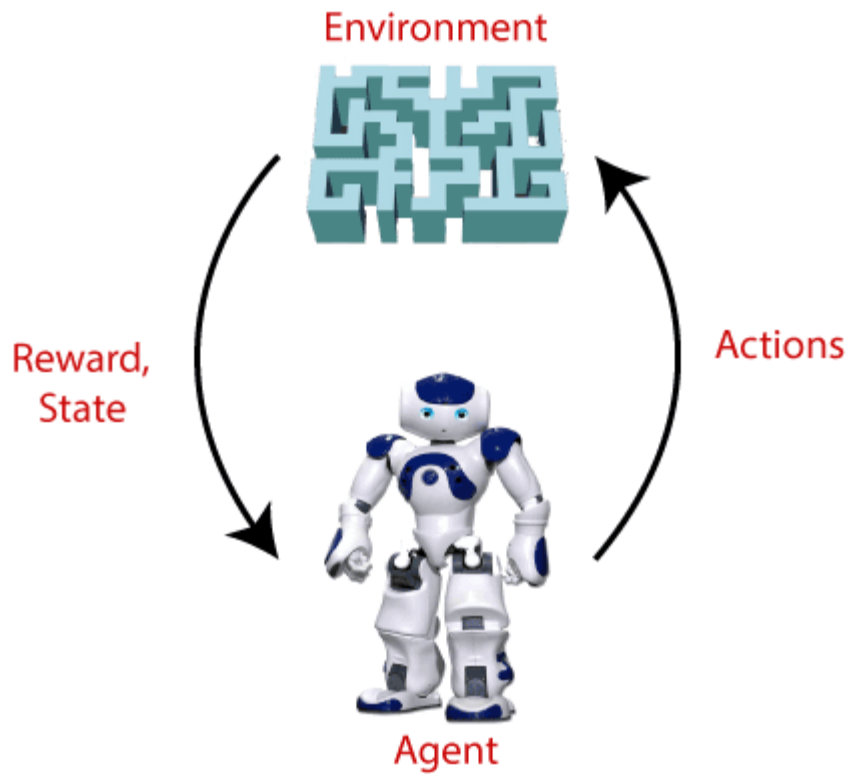
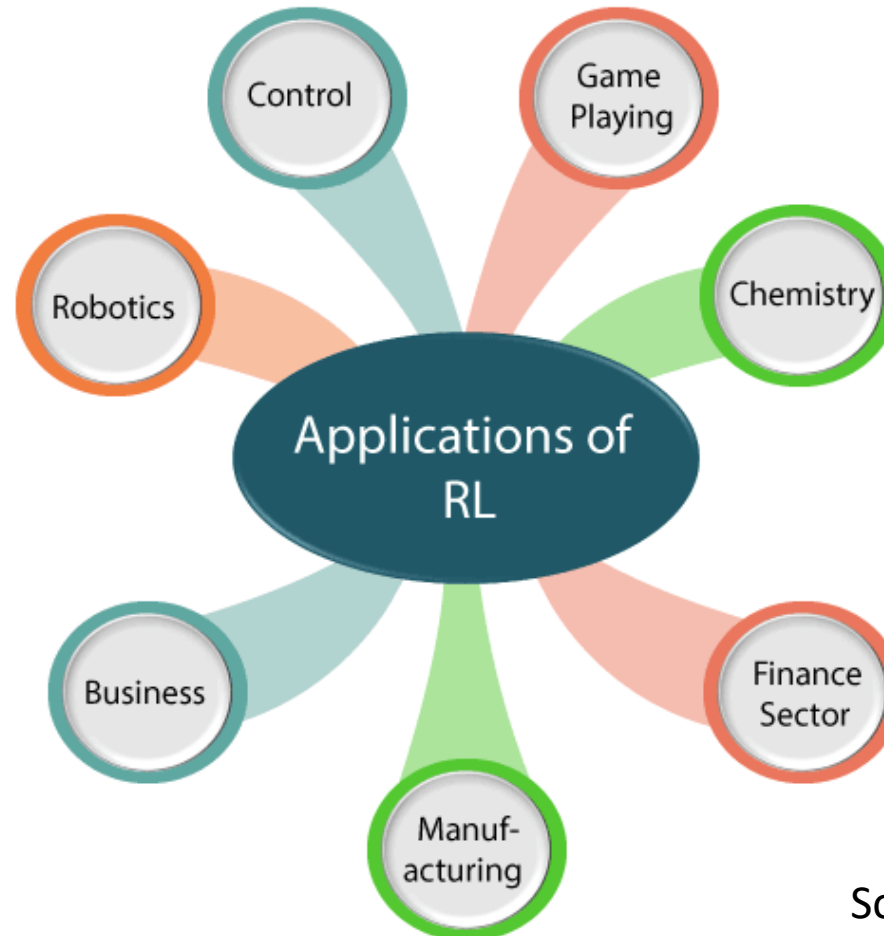Teaching a chicken by feeding

Reinforcement Learning Model
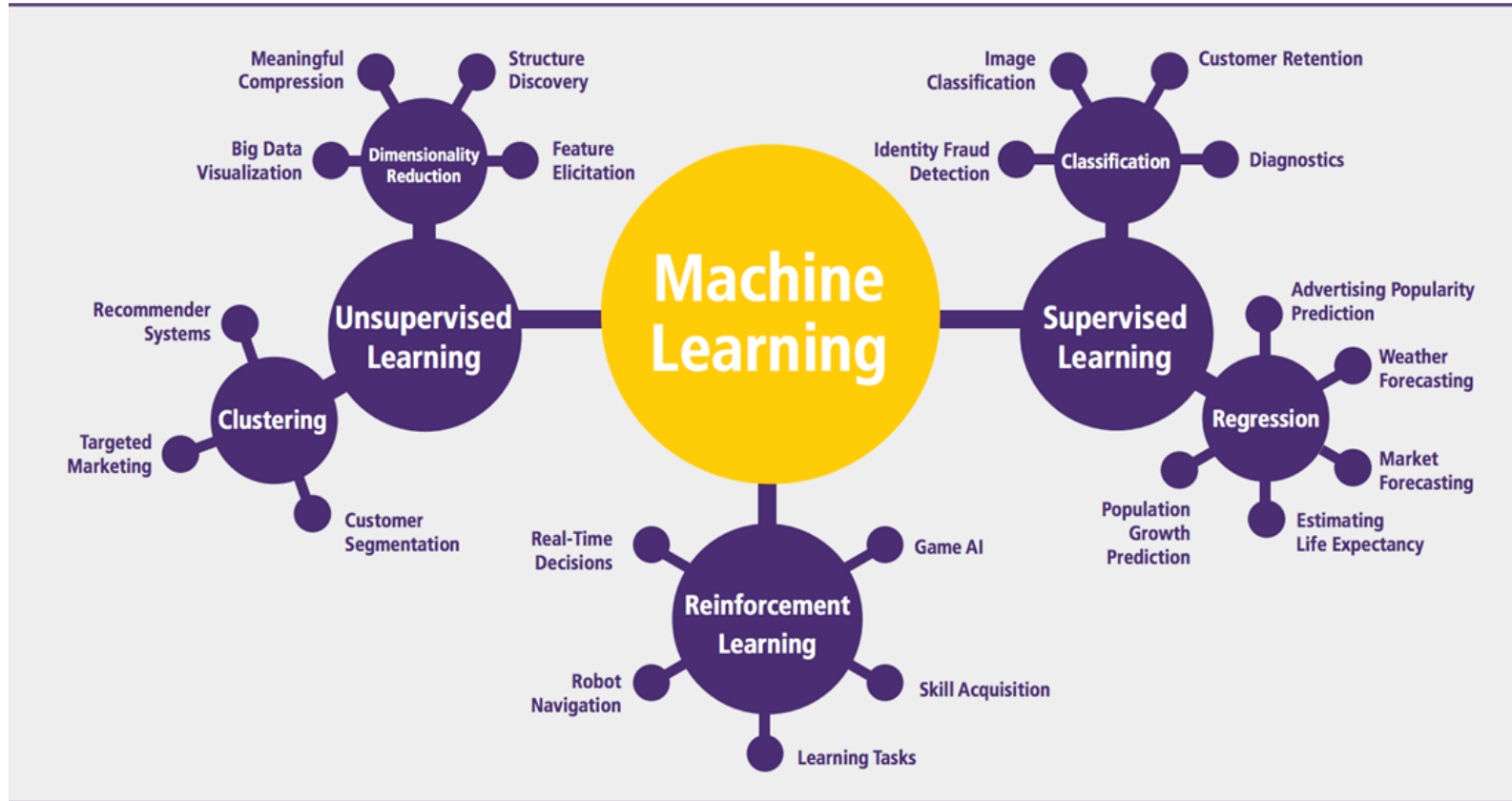
Applications of Reinforcement Learning (RL)

- Self-driving cars
- Robots
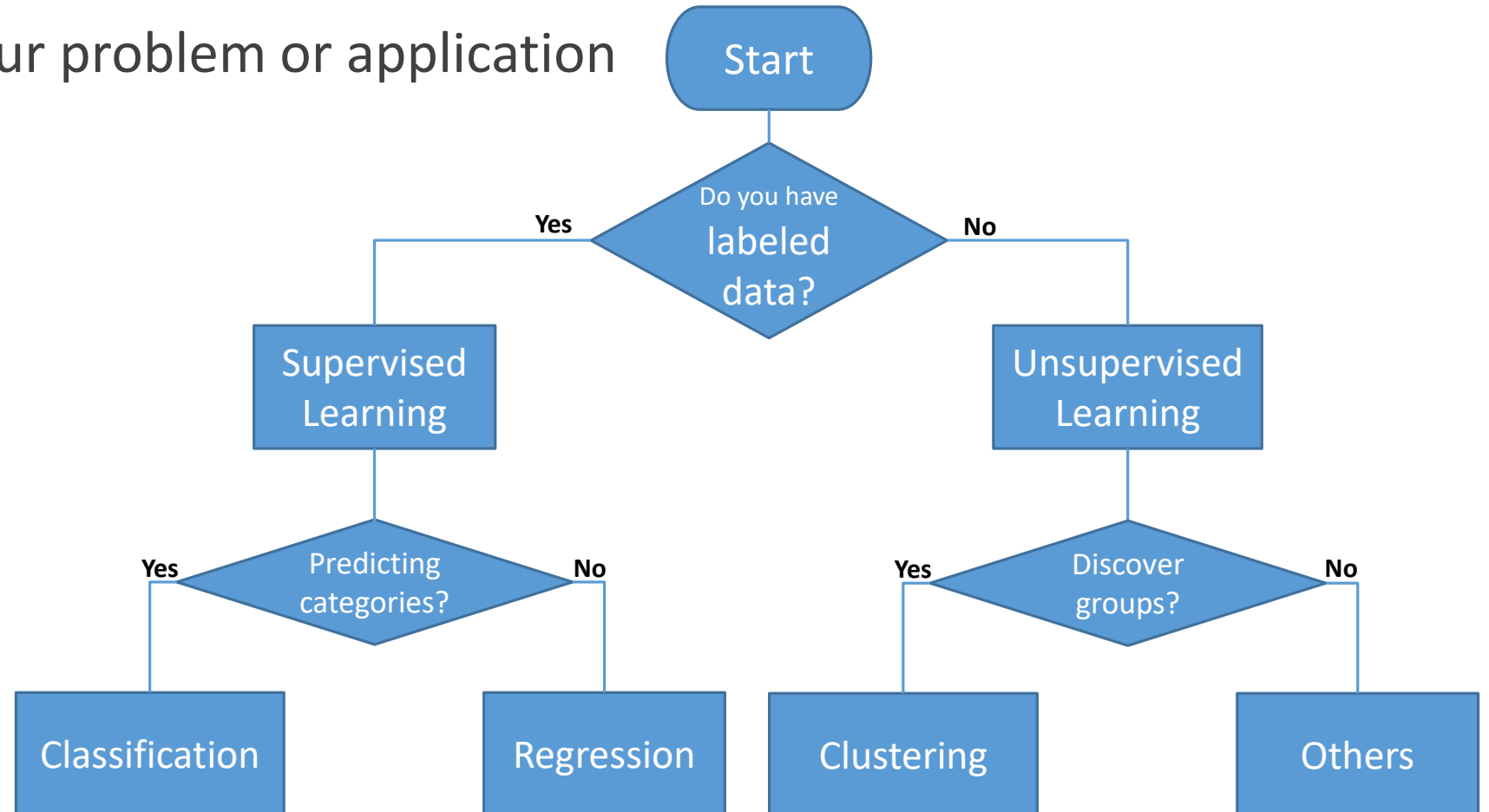- Games
- Automating trading
- Enterprise resource management

Source: javaTpoint

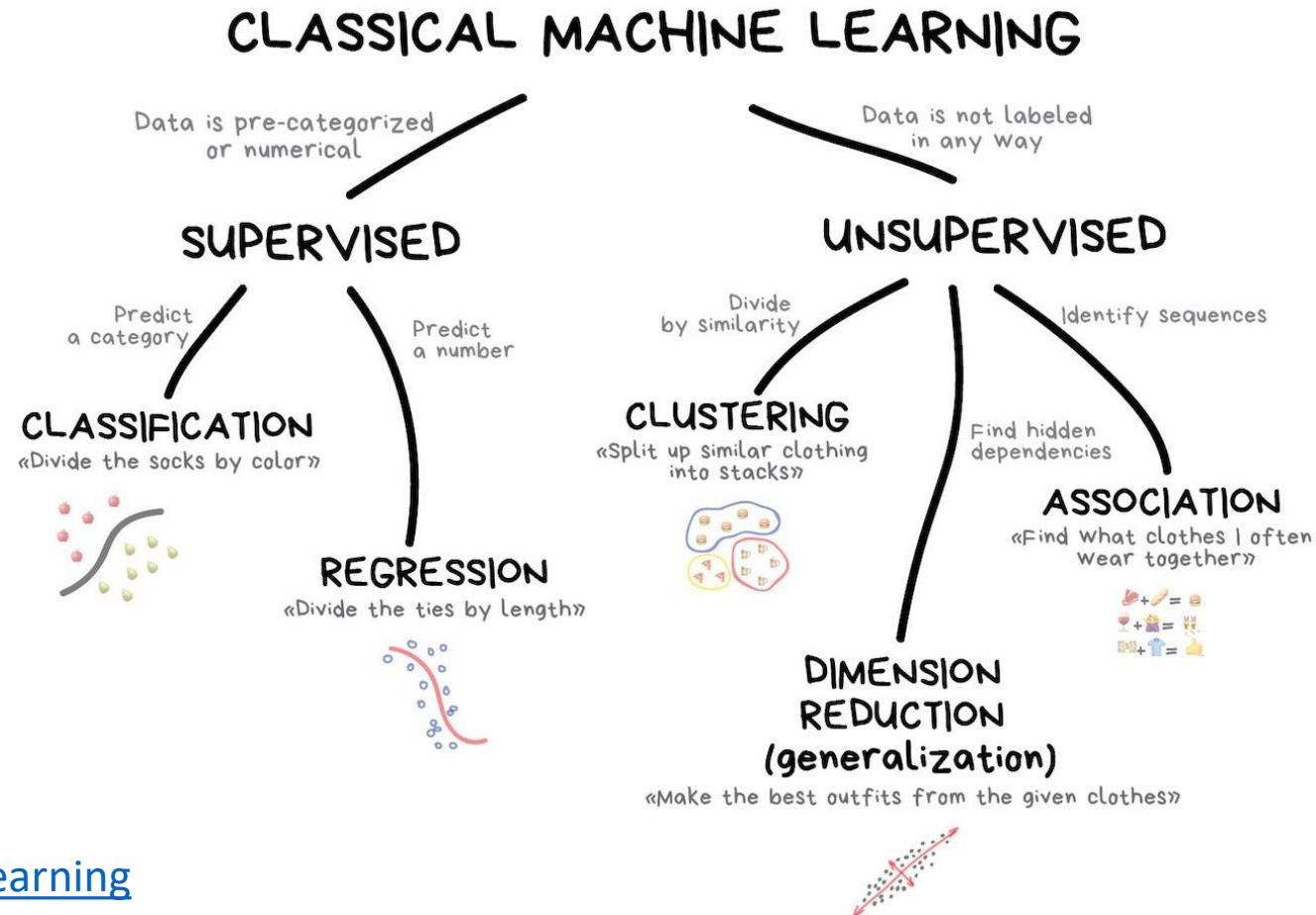# Machine Learning methods where they are used?

Which type of Machine Learning algorithm should I use?

- It depends on your problem or application

Which type of Machine Learning algorithm should I use? (Another Figure)

Source: Machine Learning

Some Learning Materials

1. Commonly used Machine Learning Algorithms [AnalyticsVidhya]

2. Machine Learning Tutorial, types of ML [javaTpoint]

3. Supervised Machine Learning [javaTpoint] [Guru99]

4. Unsupervised Machine Learning: [javaTpoint] [Guru99]

5. Reinforcement Learning Tutorial [javaTpoint]

6. Introduction to Machine Learning with scikit-learn [Kevin Markham]

7. Machine Learning for Everyone [vas3k]

8. A Machine Learning Tutorial With Examples: An Introduction to ML Theory and Its Applications, Toptal [Link]

End of Lecture-2,3