01 Introduction To Machine Learning

Dr. Nuhman Ul Haq

CUI Abbottabad

March 5, 2025

1. Machine Learning is *"the study that gives computers the ability to learn without being explicitly programmed"*

   Arthur Lee Samuel (December 5, 1901 – July 29, 1990) was an American pioneer in the field of computer gaming and artificial intelligence. He popularized the term "machine learning" in 1959. The Samuel Checkers-playing Program was among the world's first successful self-learning programs, and as such a very early demonstration of the fundamental concept of artificial intelligence (AI).

2. *"A program can be said to learn from experience E with respect to some class of task T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E"* (Tom Mitchel)

Task(T)	**Experience(E) ((training data)**	**Performance Measure (P) (Accuracy)**
Identifying cats and dog	Labeled images of cats and dogs	ratio of correctly recognized images
Recognizing hand-written digits	Labeled images of hand written digits	ratio of correctly recognized images

1. Labeled images means humans assigned correct labels to all images. **Supervised Learning**

Definition of Machine Learning?

1. "Machine learning is the training of a model from data that generalizes a decision against a performance measure." (Jason Brownlee)
2. "A branch of artificial intelligence in which a computer generates rules underlying or based on raw data that has been fed into it." (Dictionary.com)
3. "Machine learning is a scientific discipline that is concerned with the design and development of algorithms that allow computers to evolve behaviors based on empirical data, such as from sensor data or databases." (Wikipedia)

Generalization Analogy

	class	Result Student 1	Result Student 2
Calculus problems (dataset)	Problems Solved in class (training set)	9/10	9/10
	Problems not Solved in class	1/10 No Generalization	9/10 (Generalization)
Learning Task	**Training**	**Testing**	

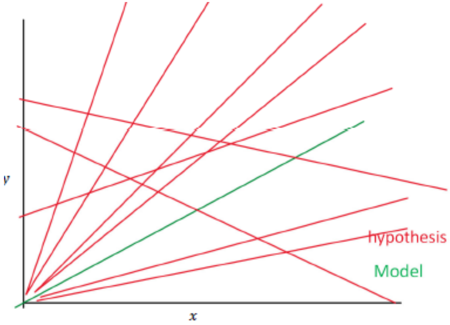
Example:Self Driving Car (Google)

Application of Machine Learning

1. Image Processing, Computer Vision and Robotics
    1. Face Recognition
    2. Hand Writing Recognition
    3. Autonomous cars
2. Business
    1. Fraud Detection
    2. Credit Approval
    3. Product Recommendation
3. Communication
    1. Spam filtering(e-mail)
    2. Analysis of caller information
4. Other Applications
    1. Recommender Systems
    2. Natural Language Processing
    3. Weather Forecasting
    4. Computer Games

Some Basic Concepts

1. **Model:** A machine learning model can be a mathematical representation of a real-world process. To generate a machine learning model, training data needs to be provided to a machine learning algorithm to learn from.
2. **Hypothesis Set:** A hypothesis set is taken at the beginning before the training start. For example, in linear regression problem, the hypothesis set consists of all linear equations (i.e. set of functions) in the plane (or n-dimensional space).
3. **Algorithm:** The machine learning algorithm chooses $t_j = h_j$ function form the hypothesis set that best fits the training data.



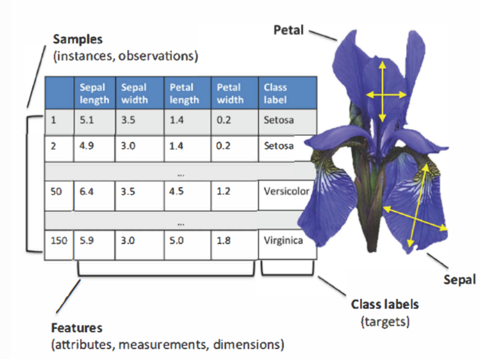
1. $y = w_0 + w_1 x$
   - Learning Parameters
   - $w_1 = Slope$
   - $w_0 = Intercept$
2. machine learning algorithm Searches to find the model from the hypotheses. Model approximately represents the relationship between the input and output.

Basic Concepts

1. **Feature:** Features are individual independent variables that act as the input to the our system (alias model). Prediction models use features to make predictions. New features can also be obtained from old features using a method known as 'feature engineering'. More simply, you can consider one column of your data set to be one feature. Sometimes these are also called attributes. And the number of features are called dimensions.
2. **Training:** While training, the training data is passed as input to the learning algorithm. The learning algorithm finds patterns in the training data such that the input parameters correspond to the target. The output of the training process is a machine learning model which you can then use to make predictions. This process is also called "learning".

Features

The Iris dataset contains the measurements of 150 Iris flowers from three different species - Setosa, Versicolor, and Virginica. Here, each flower exam pie represents one row in the dataset. Four measurements of each flower (in centimeters) are stored in each row. Thus each column 1s a feature in the dataset.



Machine Learning Sub-Fields

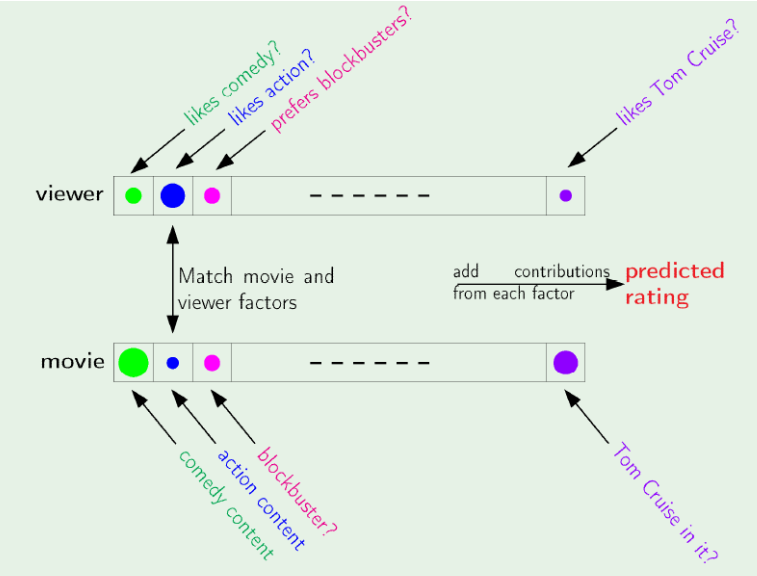
1. Supervised Learning
2. Unsupervised Learning
3. Reinforcement Learning
4. Semi-supervised Learning
5. Deep Learning and Neural Networks

Machine Learning Sub-Fields

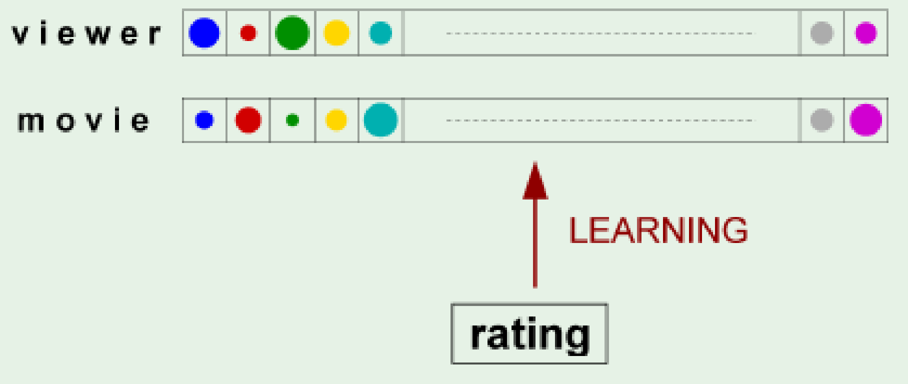
$$\textbf{Supervised Learning} = \begin{cases} \text{Labeled Data} \\ \text{Direct Feedback} \\ \text{Predict outcome/future.} \end{cases}$$

$$\textbf{Unsupervised Learning} = \begin{cases} \text{No Label} \\ \text{No Feedback} \\ \text{Find Hidden Structures.} \end{cases}$$

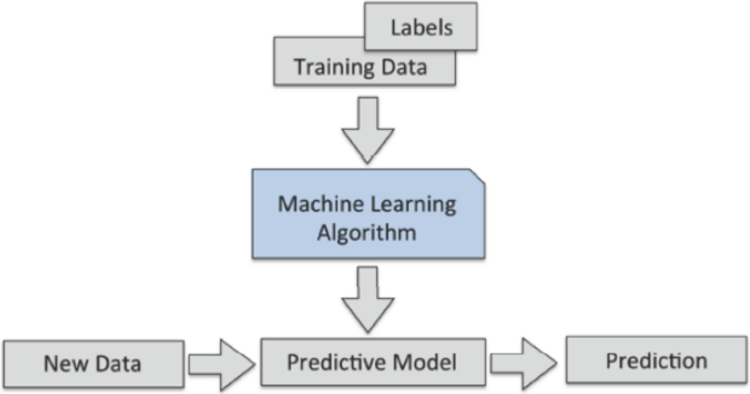
$$\textbf{Reinforcement Learning} = \begin{cases} \text{Decision Process} \\ \text{Reward System} \\ \text{Learn Series of Actions.} \end{cases}$$


likes comedy?
likes action?
prefers blockbusters?
likes Tom Cruise?
viewer
Match movie and viewer factors
add contributions from each factor
predicted rating
movie
comedy content
action content
blockbuster?
Tom Cruise in it?

Example: Movies Rating



Supervised Learning

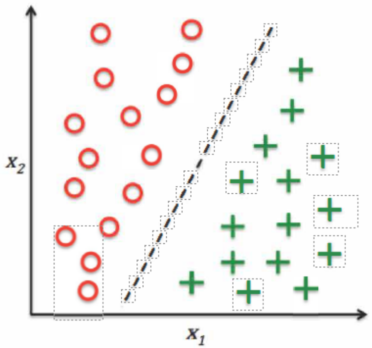


Problem Categories

1. Classification
2. Regression
3. Clustering
4. Optimization
5. Dimensionality Reduction

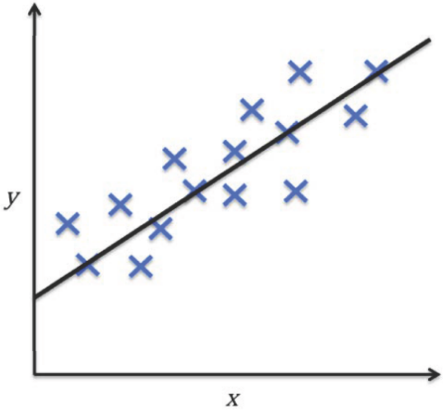
Classification Problem

A **classification problem** is a type of **supervised machine learning** task where the goal is to predict the category or class of given input data. The algorithm learns from labeled training data and then assigns labels (or categories) to new, unseen data points.



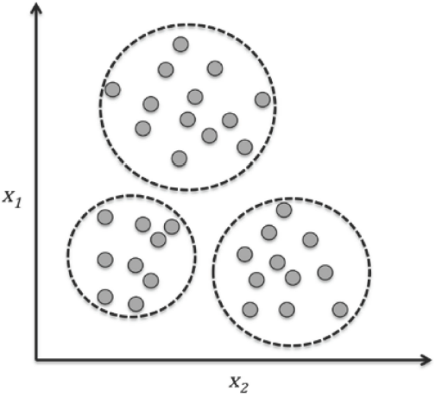
Regression Problem

A **regression problem** is a type of **supervised machine learning** task where the goal is to predict a continuous numerical value based on input features. Unlike classification problems, which predict categories, regression focuses on estimating real-valued outputs.



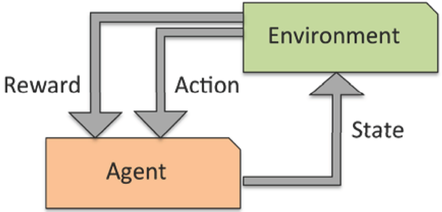
Unsupervised Learning: Clustering

Clustering is an **unsupervised machine learning** technique used to group similar data points into clusters based on their features. Unlike supervised learning, clustering does not rely on labeled data; instead, it identifies patterns or structures within the data.

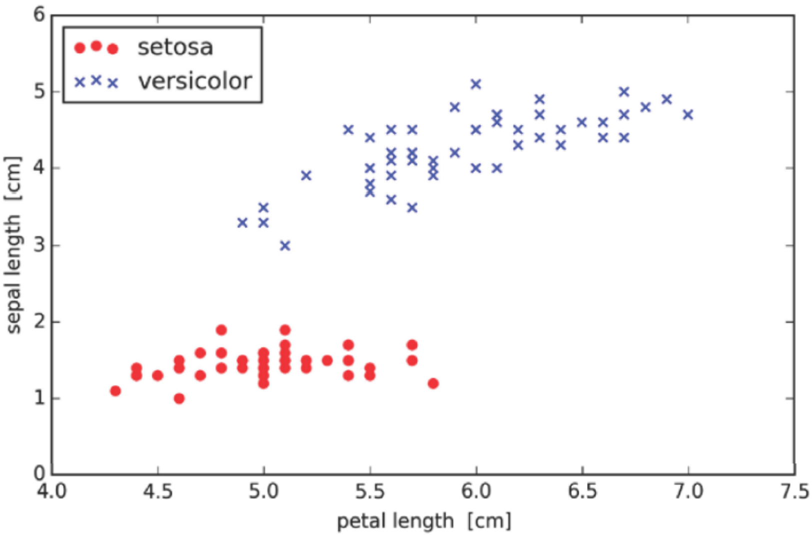


Reinforcement Learning

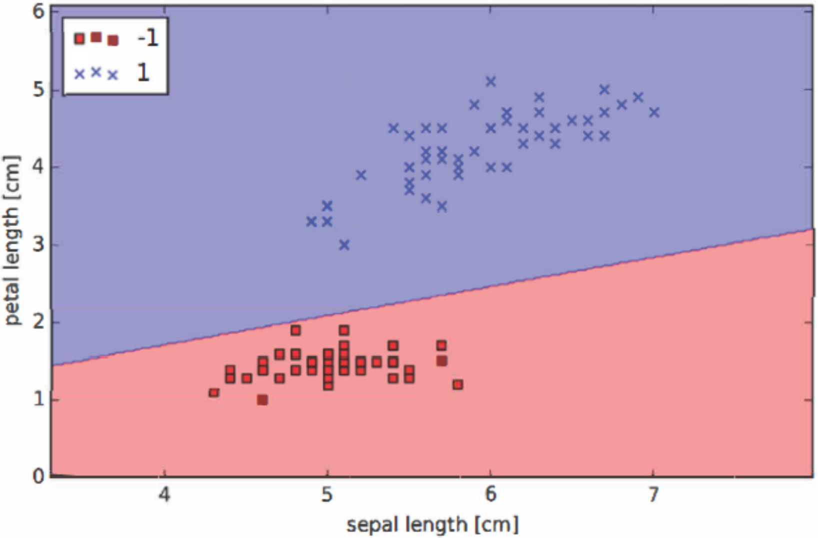
Reinforcement Learning (RL) is a type of machine learning where an agent learns to make decisions by interacting with an environment to achieve a specific goal. The agent receives feedback in the form of rewards or penalties based on its actions, and its objective is to maximize the cumulative reward over time.



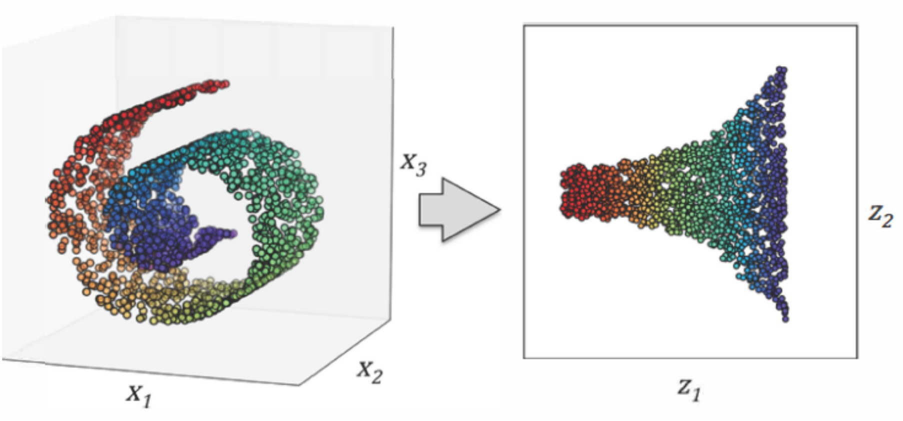
Supervised Learning (Scatterplot)



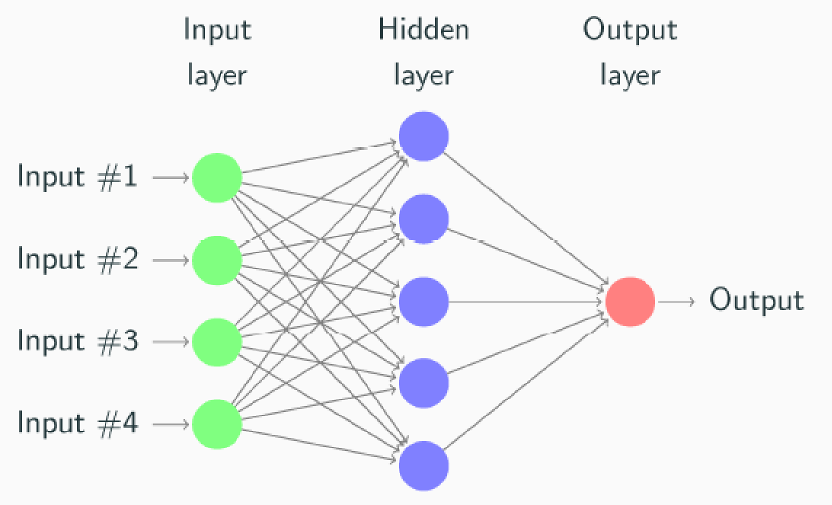
Unsupervised Learning (Scatterplot)



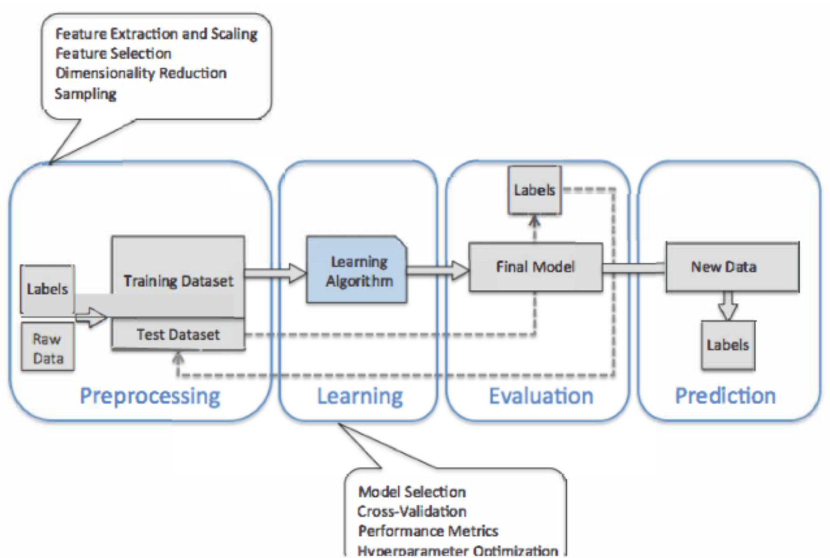
Dimensionality Reduction



Neural Networks



Machine Learning Process



Matrix Notations

1. In machine learning (mostly), each training example is represented by a row vector (a matrix having only one row) which is called feature vector. Thus a training example X having n features is given by:
   $X = x_1, x_2, ..., x_n$
2. Since there are many such training examples therefore a particular training example is represented by a superscript in parenthesis. thus an example m is written as:
   $X^{(m)} = x_1^{(m)}, x_2^{(m)}, x_3^{(m)}, ..., x_n^{(m)}$

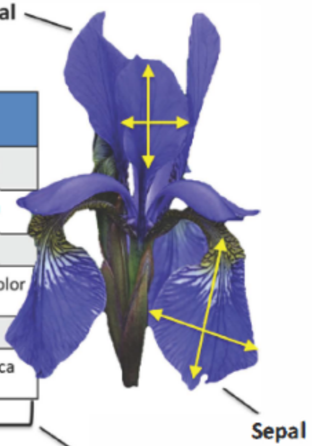
Matrix Notations

1. It is convenient to represent the entire training set as a single m x n matrix. In this case the output labels are represented as a single columns vector (matrix having only one column)

Features in the Iris dataset

Samples
(instances, observations)

	Sepal length	Sepal width	Petal length	Petal width	Class label
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
			...		
50	6.4	3.5	4.5	1.2	Versicolor
			...		
150	5.9	3.0	5.0	1.8	Virginica



Features
(attributes, measurements, dimensions)

Class labels
(targets)

Example dataset representation in Matrix

The Iris dataset, consists of 150 samples and 4 features. This can be then be written as a 150×4 matrix $X \in \mathbb{R}^{150\times 4}$

$$\begin{bmatrix} x_1^{(1)} & x_2^{(1)} & x_3^{(1)} & x_4^{(1)} \\ x_1^{(2)} & x_2^{(2)} & x_3^{(2)} & x_4^{(2)} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ x_1^{(150)} & x_2^{(150)} & x_3^{(150)} & x_4^{(150)} \end{bmatrix}$$

The labels are represented as:

$$\begin{bmatrix} y^{(1)} \\ y^{(2)} \\ . \\ . \\ . \\ y^{(150)} \end{bmatrix}$$