Artificial Intelligence

Compiled by :

Ankit Bhattarai,
Cosmos College of Management & Technology

Chapter 5

Chapter 5 (Topics to be discussed)

- Machine Learning
- Types of Learning
- Supervised Learning : Classification (Naïve Bayes/ Logistic Regression)
- Unsupervised Learning : Clustering (K-means Clustering)
- Reinforcement Learning
- Fuzzy Learning
- Boltzmann Machine
- Deep Learning

Machine Learning

- Machine learning is a type of artificial intelligence (AI) that provides computers with the ability to learn without being explicitly programmed.
- In other words, Machine learning is a branch of science that deals with programming the systems in such a way that they automatically learn and improve with experience. Here, learning means recognizing and understanding the input data and making wise decisions based on the supplied data.
- Applications of machine learning can be seen in many fields like Data mining and knowledge discovery, Speech/image/video (pattern) recognition, Adaptive control, Autonomous vehicles/robots, Decision support systems, Bioinformatics

Example

Facebook's News Feed uses machine learning to personalize each member's feed. If a member frequently stops scrolling in order to read or "like" a particular friend's posts, the News Feed will start to show more of that friend's activity earlier in the feed.

Behind the scenes, the software is simply using statistical analysis and predictive analytics to identify patterns in the user's data and use to patterns to populate the News Feed.

Types of Learning

- Rote Learning, Direct instruction (by being told), Learning by analogy, Inductive Learning and Learning by deduction

1. **Rote Learning**

- Rote learning is a technique which focuses on memorization
- It avoids understanding the inner complexities and inferences of the subject that is being learned and instead focuses on memorizing the material
- Example:- Memorizing multiplication tables, formulate , etc.

Types of Learning

2. Direct instruction

- This type of learning requires more inference than rote learning since the knowledge must be transformed into an operational form before learning when a teacher presents a number of facts directly to us in a well-organized manner.
- It is more complex than rote learning.

3. Learning by analogy

- Analogical learning is the process of learning a new concept or solution through the use of similar known concepts or solutions

- We use this type of learning when solving problems on an exam where previously learned examples serve as a guide or when make frequent use of analogical learning.

4. Inductive Learning (Learning by examples)

- Based on classification of problems prior to solving
- It is a powerful form of learning like analogical learning which also requires more inferring than the first two methods. This learning requires the use of inductive inference, a form of invalid but useful inference.
- Based on induction of result
- For example we learn the concepts of color or sweet taste after experiencing the sensations associated with several examples of colored objects or sweet foods

5. Learning by deduction

- Deductive learning is accomplished through a sequence of deductive inference steps using known facts.
- From the known facts, new facts or relationships are logically derived.
- Deductive learning usually requires more inference than the other methods.

Explanation based learning (EBL)

- Explanation-based learning (EBL) is a form of machine learning that exploits a very strong, or even perfect, domain theory to make generalizations or form concepts from training examples.
- An Explanation-based Learning (EBL) system accepts an example (i.e. a training example) and explains what it learns from the example.
- The EBL system takes only the relevant aspects of the training. This explanation is translated into particular form that a problem solving program can understand. The explanation is generalized so that it can be used to solve other problems.

Explanation based learning (EBL)

- An example of EBL using a perfect domain theory is a program that learns to play chess by being shown examples.
- A specific chess position that contains an important feature, say, "Forced loss of black queen in two moves," includes many unrelated features, such as the specific scattering of pawns on the board.
- EBL can take a single training example and determine what the relevant features are in order to form a generalization

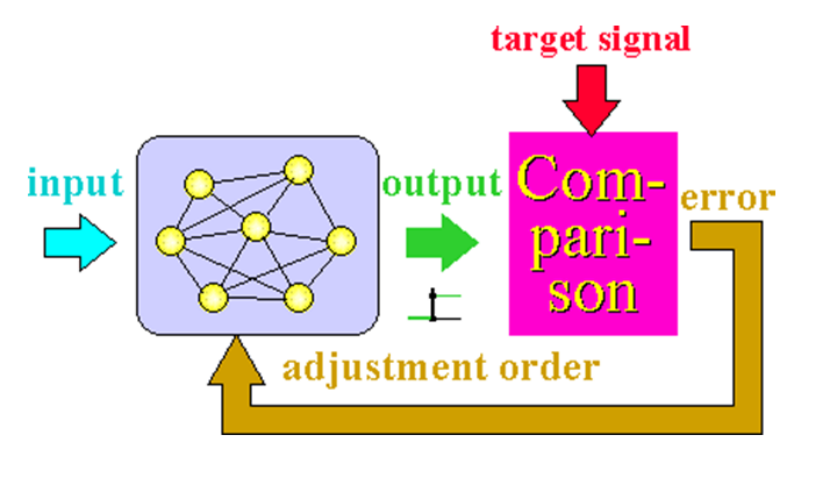
Explanation based learning (EBL)

- An EBL accepts 4 kinds of input:
  - ✓ A training example: What the learning sees in the world.
  - ✓ A goal concept: A high level description of what the program is supposed to learn.
  - ✓ A operational criterion: A description of which concepts are usable.
  - ✓ A domain theory: A set of rules that describe relationships between objects and actions in a domain.

Supervised Learning

- The network is presented with inputs together with the target (teacher signal) outputs.
- The neural network tries to produce an output as close as possible to the target signal by adjusting the values of internal weights.
- The most common supervised learning method is the “error correction method”.
- Error correction method is used for networks which their neurons have discrete output functions. Neural networks are trained with this method in order to reduce the error (difference between the network's output and the desired output) to zero.

Supervised Learning



- **Example**: For instance, suppose you are given a basket filled with different kinds of fruits. Now the first step is to train the machine with all different fruits one by one like this: Apple, Banana, Oranges, etc.

Supervised Learning

- If shape of object is rounded and depression at top having color Red then it will be labeled as –**Apple**.
- Now suppose after training the data, you have given a new separate fruit say Banana from basket and asked to identify it.
- Since machine has already learnt the things from previous data and this time have to use it wisely. It will first classify the fruit with its shape and color, and would confirm the fruit name as BANANA and put it in Banana category. Thus machine learns the things from training data(basket containing fruits) and then apply the knowledge to test data(new fruit).

Supervised Learning: Classification

- Supervised learning classified into two categories of algorithms:

I. **Classification:** A classification problem is when the output variable is a category, such as “Red” or “blue” or “disease” and “no disease”. A classification model attempts to draw some conclusion from observed values. Given one or more inputs a classification model will try to predict the value of one or more outcomes.

Example:

A common example of classification comes with detecting spam emails. To write a program to filter out spam emails, a computer programmer can train a machine learning algorithm with a set of spam-like emails labelled as spam and regular emails labelled as not-spam.

Supervised Learning: Classification (Naïve Bayes)

- Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems.
- In machine learning, naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes theorem with strong (naive) independence assumptions between the features.
- Naive Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem.
- Some popular examples of Naïve Bayes Algorithm are **spam filtration, Sentimental analysis, and classifying articles**.

Supervised Learning: Classification (Naïve Bayes)

The Naïve Bayes algorithm is comprised of two words Naïve and Bayes, Which can be described as:

- **Naïve**: It is called Naïve because it assumes that the occurrence of a certain feature is independent of the occurrence of other features. Such as if the fruit is identified on the bases of color, shape, and taste, then red, spherical, and sweet fruit is recognized as an apple. Hence each feature individually contributes to identify that it is an apple without depending on each other.
- **Bayes**: It is called Bayes because it depends on the principle of Bayes' Theorem.

Supervised Learning: Classification (Naïve Bayes)

- Bayes' theorem is also known as **Bayes' Rule** or **Bayes' law**, which is used to determine the probability of a hypothesis with prior knowledge. It depends on the conditional probability.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Where,

P(A|B) is Posterior probability: Probability of hypothesis A on the observed event B.

P(B|A) is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true.

P(A) is Prior Probability: Probability of hypothesis before observing the evidence.

P(B) is Marginal Probability: Probability of Evidence.

Supervised Learning: Classification (Naïve Bayes)

Advantages

1. Naïve Bayes is one of the fast and easy ML algorithms to predict a class of datasets.
2. It can be used for Binary as well as Multi-class Classifications.
3. It performs well in Multi-class predictions as compared to the other Algorithms.
4. It is the most popular choice for **text classification problems**.

Disadvantages

1. Naive Bayes assumes that all features are independent or unrelated, so it cannot learn the relationship between features.

Numerical on Naïve Bayes

Supervised Learning: Classification (K Nearest Neighbor)

- K Nearest Neighbor is a simple algorithm that stores all the available cases and classifies the new data or case based on a similarity measure.
- K-Nearest Neighbor is one of the simplest Machine Learning algorithms based on Supervised Learning technique.
- K-NN is a **non-parametric algorithm**, which means it does not make any assumption on underlying data.
- K-NN is also called a **lazy learner algorithm** because it does not learn from the training set immediately instead it stores the dataset and at the time of classification, it performs an action on the dataset.

Supervised Learning: Classification (K Nearest Neighbor)

- **Example:** Suppose, we have an image of a creature that looks similar to cat and dog, but we want to know either it is a cat or dog.
- So for this identification, we can use the KNN algorithm, as it works on a similarity measure.
- Our KNN model will find the similar features of the new data set to the cats and dogs images and based on the most similar features it will put it in either cat or dog category.

Supervised Learning: Classification (K Nearest Neighbor)

The K-NN working can be explained on the basis of the below algorithm:

Step-1: Select the number K of the neighbors

Step-2: Calculate the Euclidean distance of **K number of neighbors**

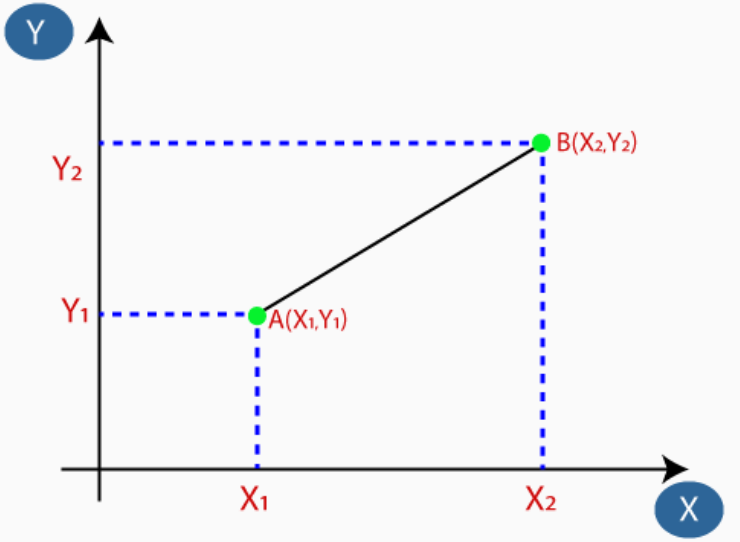
Step-3: Take the K nearest neighbors as per the calculated Euclidean distance.

Step-4: Among these k neighbors, count the number of the data points in each category.

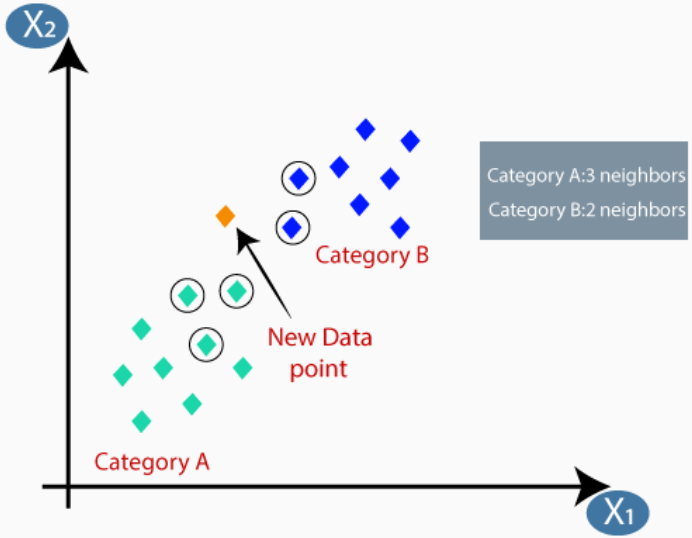
Step-5: Assign the new data points to that category for which the number of the neighbor is maximum.

Step-6: Our model is developed.

Supervised Learning: Classification (K Nearest Neighbor)



Euclidean Distance between A_1 and $B_2 = \sqrt{(X_2-X_1)^2+(Y_2-Y_1)^2}$



As we can see the 3 nearest neighbors are from category A, hence this new data point must belong to category A.

Supervised Learning: Classification (K Nearest Neighbor)

Pros of KNN

1.Simple to implement

2.Flexible to feature/distance choices

3.Naturally handles multi-class cases

4.Can do well in practice with enough representative data

Cons of KNN

1.Need to determine the value of parameter K (number of nearest neighbors)

2.Computation cost is quite high because we need to compute the distance of each query instance to all training samples.

3.Storage of data

Supervised Learning: Regression

II. Regression: A regression problem is when the output variable is a real or continuous value, such as “salary” or “weight”. Many different models can be used, the simplest is the linear regression. It tries to fit data with the best hyper-plane which goes through the points**.** There are three types of regression: Linear Regression, Logistic Regression & Ridge Regression.

Example:

For example, a house may be predicted to sell for a specific dollar value, perhaps in the range of $100,000 to $200,000.

Supervised Learning : Regression (Logistic Regression)

- Logistic regression is used whenever you need to assign an input to one of several classes.
- It applies a logistic function to a linear combination of features, yielding a probability for an input to be in one of the classes.
- The output is usually binary, but logistic regression can also be applied to multiclass classification problems.
- Logistic regression is used in various fields, including machine learning, most medical fields, and social sciences.

Supervised Learning : Regression (Logistic Regression)

- Logistic regression is used whenever you need to assign an input to one of several classes.
- It applies a logistic function to a linear combination of features, yielding a probability for an input to be in one of the classes.
- The output is usually binary, but logistic regression can also be applied to multiclass classification problems.
- Logistic regression is used in various fields, including machine learning, most medical fields, and social sciences.

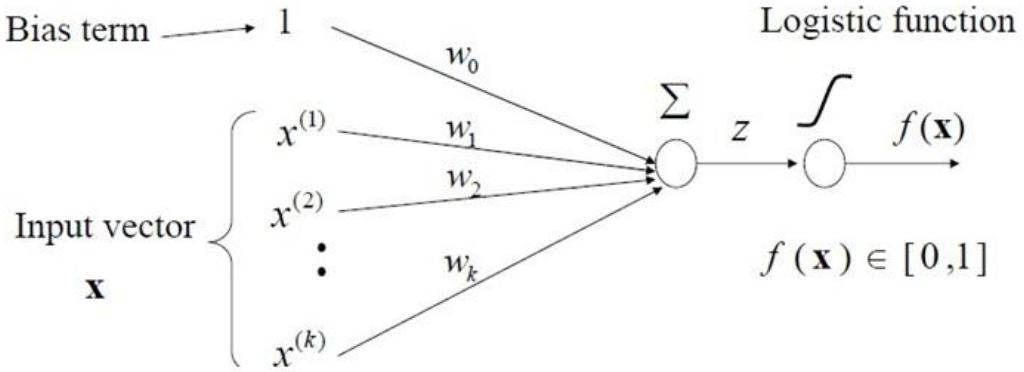
Supervised Learning : Regression (Logistic Regression)

- A function model with smooth switching:

$$f(\mathbf{x}) = g(w_0 + w_1 x^{(1)} + \ldots w_k x^{(k)})$$

where $\boldsymbol{w}$ are parameters of the models

and $g(z)$ is a **logistic function** $g(z) = 1/(1+e^{-z})$



Unsupervised Learning

- Unsupervised learning is the training of machine using information that is neither classified nor labeled and allowing the algorithm to act on that information without guidance.
- In unsupervised learning, there is no teacher (target signal) from outside and the network adjusts its weights in response to only the input patterns.
- The goal for unsupervised learning is to model the underlying structure or distribution in the data in order to learn more about the data.
- ***Example***: For instance, suppose it is given an image having both dogs and cats which have not seen ever.

Unsupervised Learning

- Thus machine has no any idea about the features of dogs and cat so we can't categorize it in dogs and cats. But it can categorize them according to their similarities, patterns and differences i.e., we can easily categorize the above picture into two parts.
- First may contain all pics having **dogs** in it and second part may contain all pics having **cats** in it. Here you didn't learn anything before, means no training data or examples.
- Unsupervised learning classified into two categories of algorithms:

i. Clustering

ii. Association

Clustering:

- Cluster analysis or clustering is the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar to each other than to those in other groups (clusters). It is a main task of exploratory data mining, and a common technique for statistical data analysis, used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, bioinformatics and data compression.
- Cluster analysis itself is not one specific algorithm, but the general task to be solved. It can be achieved by various algorithms that differ significantly in their notion of what constitutes a cluster and how to efficiently find them. Popular notions of clusters include groups with small distances among the cluster members, dense areas of the data space, intervals or particular statistical distributions. Clustering can therefore be formulated as a multi-objective optimization problem.

Unsupervised Learning : Clustering (K-means Clustering)

- K-means clustering is a type of unsupervised learning, which is used when you have unlabeled data (i.e., data without defined categories or groups**).**
- We are given a data set of items, with certain features, and values for these features (like a vector). The task is to categorize those items into groups. To achieve this, we will use the K Means algorithm; an unsupervised learning algorithm.
- A cluster refers to a collection of data points aggregated together because of certain similarities.
- K means algorithm is very popular and used in a variety of applications such as market segmentation, document clustering, image segmentation and image compression, etc.

Unsupervised Learning : Clustering (K-means Clustering)

The **algorithm** works as follows:

i. First we initialize k points, called means, randomly.

ii. We categorize each item to its closest mean and we update the mean's coordinates, which are the averages of the items categorized in that mean so far.

iii. We repeat the process for a given number of iterations and at the end, we have our clusters.

Unsupervised Learning : Clustering (K-means Clustering)

- For example, in computer graphics, color quantization is the task of reducing the color palette of an image to a fixed number of colors k. The k-means algorithm can easily be used for this task and produces competitive results. A use case for this approach is image segmentation. Other uses of vector quantization include non-random sampling, as k-means can easily be used to choose k different but prototypical objects from a large data set for further analysis.
- It often is used as a preprocessing step for other algorithms, for example to find a starting configuration.

Numerical on K Means Clustering

Dimensionality Reduction

In pattern recognition, Dimension Reduction is defined as-

- It is a process of converting a data set having vast dimensions into a data set with lesser dimensions.
- It ensures that the converted data set conveys similar information concisely.

Dimension reduction offers several benefits such as-

- It compresses the data and thus reduces the storage space requirements.
- It reduces the time required for computation since less dimensions require less computation.
- It eliminates the redundant features.
- It improves the model performance.

Dimensionality Reduction: Principal Components Analysis (PCA)

- An exploratory technique used to reduce the dimensionality of the data set to 2D or 3D.

- Can be used to:

✓ Reduce number of dimensions in data

✓ Find patterns in high-dimensional data

✓ Visualize data of high dimensionality

- Example applications: Face recognition, Image compression and Gene expression analysis

Dimensionality Reduction: Principal Components Analysis (PCA)

- PCA is the process of computing the principal components and using them to perform a change of basis on the data, sometimes using only the first few principal components and ignoring the rest.
- It transforms the variables into a new set of variables called as principal components.
- There can be only two principal components for a two-dimensional data set.
- PCA is used in exploratory data analysis and for making predictive models.
- It is commonly used for dimensionality reduction by projecting each data point onto only the first few principal components to obtain lower-dimensional data while preserving as much of the data's variation as possible

Dimensionality Reduction(Principal Component Analysis)

The steps involved in PCA Algorithm are as follows:

Step 01: Get data.

Step 02: Compute the mean vector (μ).

Step 03: Subtract mean from the given data.

Step 04: Calculate the covariance matrix.

Step 05: Calculate the eigen vectors and eigen values of the covariance matrix.

Step 06: Choosing components and forming a feature vector.

Step 07: Deriving the new data set.

Dimensionality Reduction (Linear Discriminant Analysis (LDA)**)**

- LDA is a dimensionality reduction algorithm, similar to PCA.
- However while PCA is an unsupervised algorithm that focusses on maximizing variance in a dataset, LDA is a supervised algorithm that maximizes separability between classes.
- Linear discriminant analysis is primarily used here to reduce the number of features to a more manageable number before classification. Each of the new dimensions is a linear combination of pixel values, which form a template.
- The original dichotomous discriminant analysis was developed by Sir Ronald Fisher in 1936

Unsupervised Learning : Dimensionality Reduction(Linear Discriminant Analysis (LDA))

- **Applications**

1. Bankruptcy prediction
2. Face recognition
3. Earth science
4. Marketing
5. Biomedical studies

Support Vector Machine

- In machine learning, support-vector machines (SVMs, also support-vector networks) are supervised learning models with associated learning algorithms that analyze data for classification and regression analysis.
- Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier.

Support Vector Machine

- **Example:**

✓ Suppose we see a strange cat that also has some features of dogs, so if we want a model that can accurately identify whether it is a cat or dog, so such a model can be created by using the SVM algorithm.

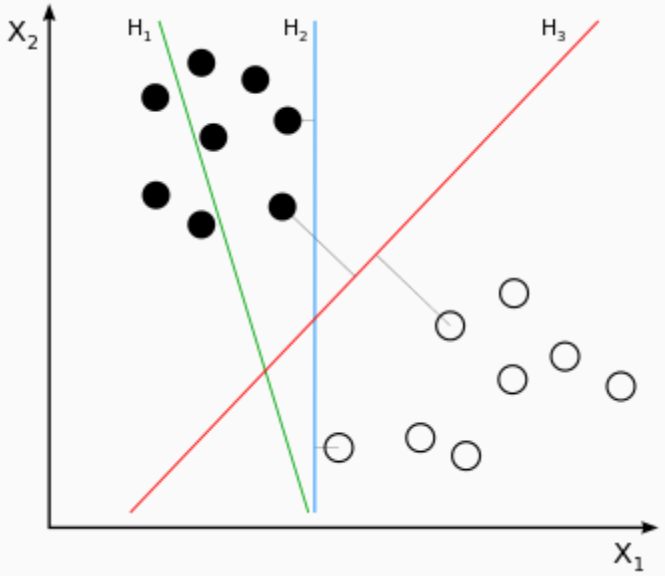
✓ We will first train our model with lots of images of cats and dogs so that it can learn about different features of cats and dogs, and then we test it with this strange creature.

✓ So as support vector creates a decision boundary between these two data (cat and dog) and choose extreme cases (support vectors), it will see the extreme case of cat and dog.

✓ On the basis of the support vectors, it will classify it as a cat.

Support Vector Machine

- The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future.
- This best decision boundary is called a hyperplane.
- Here H_3 is best hyperplane that separates them with the maximal margin.



Support Vector Machine

SVM can be of two types:

Linear SVM: Linear SVM is used for linearly separable data, which means if a dataset can be classified into two classes by using a single straight line, then such data is termed as linearly separable data, and classifier is used called as Linear SVM classifier.

Non-linear SVM: Non-Linear SVM is used for non-linearly separated data, which means if a dataset cannot be classified by using a straight line, then such data is termed as non-linear data and classifier used is called as Non-linear SVM classifier.

Support Vector Machine : Applications

- SVMs are helpful in text and hypertext categorization.
- Classification of images can also be performed using SVMs.
- Classification of satellite data like SAR data using supervised SVM.
- Hand-written characters can be recognized using SVM.

- **Semi-Supervised learning** is a type of Machine Learning algorithm that represents the intermediate ground between Supervised and Unsupervised learning algorithms. It uses the combination of labeled and unlabeled datasets during the training period.
- **Applications**: Speech analysis, web content classification, protein sequence classification and text document classification.

Semi Supervised Learning

Reinforcement Learning

- Reinforcement learning is similar to supervised learning in that some feedback is given, however instead of providing a target output a reward is given based on how well the system performed
- The aim of reinforcement learning is to maximize the reward the system receives through trial-and error.
- This paradigm relates strongly with how learning works in nature, for example an animal might remember the actions it's previously taken which helped it to find reward.

Reinforcement Learning

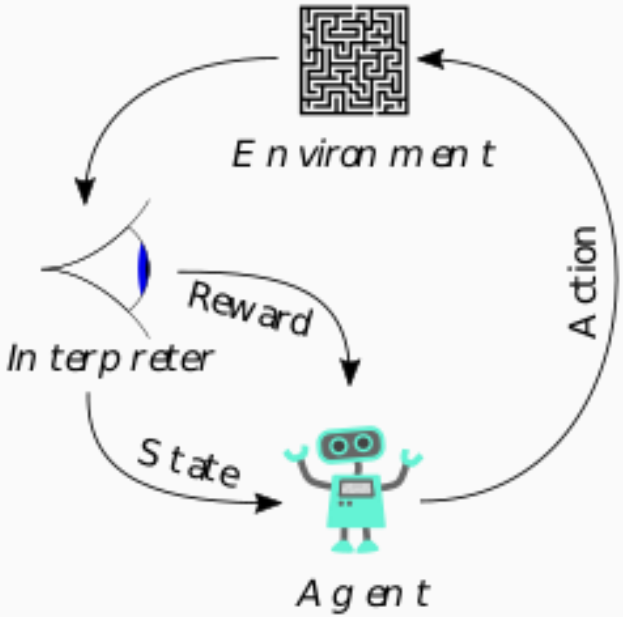


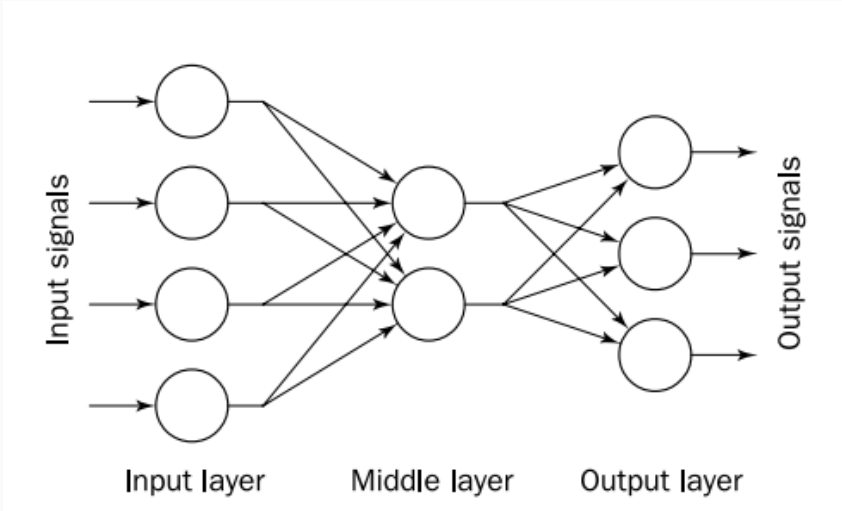
Fig. Reinforcement Learning

Applications

- Robotics for industrial automation.
- Business strategy planning
- Machine learning and data processing
- It helps you to create training systems that provide custom instruction and materials according to the requirement of students.
- Aircraft control and robot motion control

Neural Network

- A neural network or Artificial Neural Network is a network or circuit of neurons composed of artificial neurons or nodes.
- An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information.
- ANNs, like people, learn by example.
- An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process.



Fuzzy Learning

- The term fuzzy refers to things which are not clear or are vague. In the real world many times we encounter a situation when we can’t determine whether the state is true or false, their fuzzy logic provides a very valuable flexibility for reasoning. In this way, we can consider the inaccuracies and uncertainties of any situation.
- In Boolean system truth value, 1.0 represents absolute truth value and 0.0 represents absolute false value.
- In the fuzzy system, there is no logic for absolute truth and absolute false value. But in fuzzy logic, there is intermediate value too present which is partially true and partially false.

Fuzzy Learning

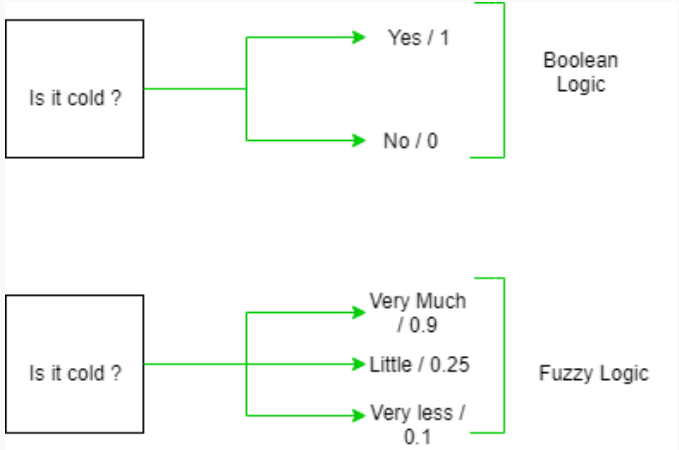


Fig. Boolean Logic Vs Fuzzy Logic

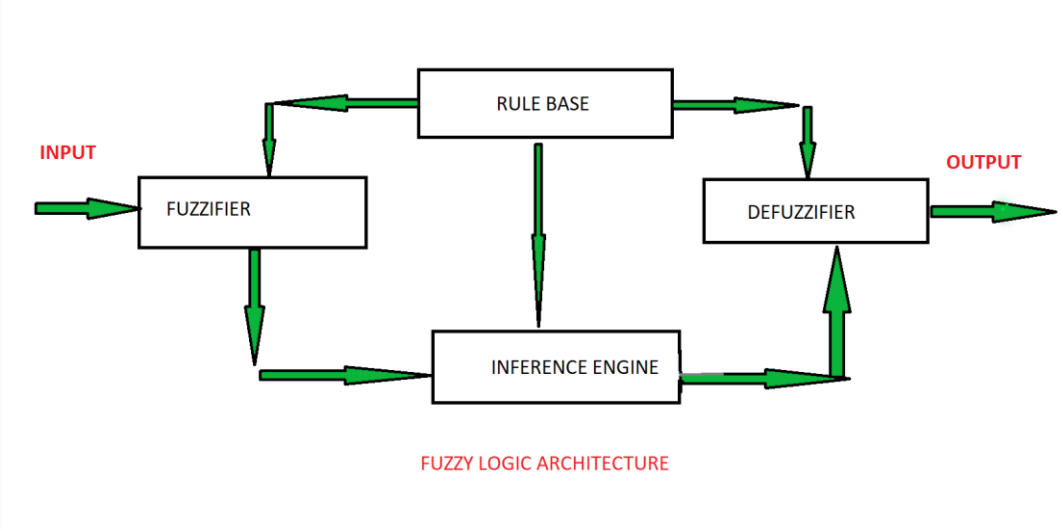


Fig. Architecture of Fuzzy Logic

Fuzzy Learning: Architecture of Fuzzy Logic

Its Architecture contains four parts:

1. **RULE BASE**: It contains the act of rules and the IF-THEN conditions provided by the experts to govern the decision making system, on the basis of linguistic information. Recent developments in fuzzy theory offer several effective methods for the design and tuning of fuzzy controllers. Most of these developments reduce the number of fuzzy rules.
2. **FUZZIFICATION:** It is used to convert inputs i.e. crisp numbers into fuzzy sets. Crisp inputs are basically the exact inputs measured by sensors and passed into the control system for processing, such as temperature, pressure, rpm's, etc.

3. INFERENCE ENGINE: It determines the matching degree of the current fuzzy input with respect to each rule and decides which rules are to be fired according to the input field. Next, the fired rules are combined to form the control actions.

4. DEFUZZIFICATION: It is used to convert the fuzzy sets obtained by inference engine into a crisp value. There are several defuzzification methods available and the best suited one is used with a specific expert system to reduce the error.

Fuzzy Learning: Advantages of Fuzzy Logic System

1. This system can work with any type of inputs whether it is imprecise, distorted or noisy input information.
2. The construction of Fuzzy Logic Systems is easy and understandable.
3. Fuzzy logic comes with mathematical concepts of set theory and the reasoning of that is quite simple.
4. It provides a very efficient solution to complex problems in all fields of life as it resembles human reasoning and decision making.
5. The algorithms can be described with little data, so little memory is required.

Fuzzy Learning: Disadvantages of Fuzzy Logic System

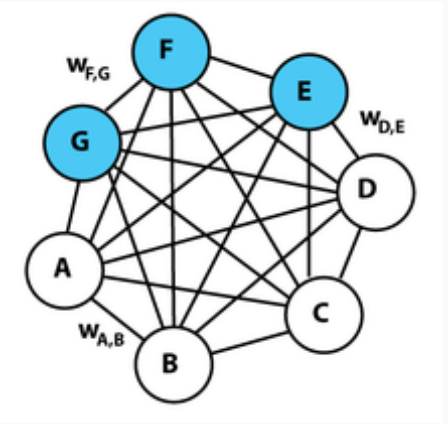
1. Many researchers proposed different ways to solve a given problem through fuzzy logic which lead to ambiguity. There is no systematic approach to solve a given problem through fuzzy logic.
2. Proof of its characteristics is difficult or impossible in most cases because every time we do not get mathematical description of our approach.
3. As fuzzy logic works on precise as well as imprecise data so most of the time accuracy is compromised.

Fuzzy Learning: Application of Fuzzy Logic System

1. It is used in the aerospace field for altitude control of spacecraft and satellite.
2. It has used in the automotive system for speed control, traffic control.
3. It is used for decision making support systems and personal evaluation in the large company business.
4. It has application in chemical industry for controlling the pH, drying, chemical distillation process.
5. Fuzzy logic are used in Natural language processing and various intensive applications in Artificial Intelligence.
6. Fuzzy logic are extensively used in modern control systems such as expert systems.
7. Fuzzy Logic is used with Neural Networks as it mimics how a person would make decisions, only much faster. It is done by Aggregation of data and changing into more meaningful data by forming partial truths as Fuzzy sets.

Boltzmann Machine

- These are stochastic learning processes having recurrent structure and are the basis of the early optimization techniques used in ANN.
- They consist of stochastic neurons, which have one of the two possible states, either 1 or 0.
- If we apply simulated annealing on discrete Hopfield network, then it would become Boltzmann Machine.
- **Objective:** The main purpose of Boltzmann Machine is to optimize the solution of a problem. It is the work of Boltzmann Machine to optimize the weights and quantity related to that particular problem.



Boltzmann Machine

- A Boltzmann machine, like a Hopfield network, is a network of units with an "energy" defined for the overall network. Its units produce binary results. Unlike Hopfield nets, Boltzmann machine units are stochastic. The global energy, E, in a Boltzmann machine is identical in form to that of a Hopfield network:

$$E = -\left(\sum_{i<j} w_{ij}\, s_i\, s_j + \sum_i \theta_i\, s_i\right)$$

Where:

- w_{ij} is the connection strength between unit j and unit i.
- s_i is the state, $s_i \in \{0, 1\}$, of unit i.
- θ_i is the bias of unit i in the global energy function. ($-\theta_i$ is the activation threshold for the unit.)

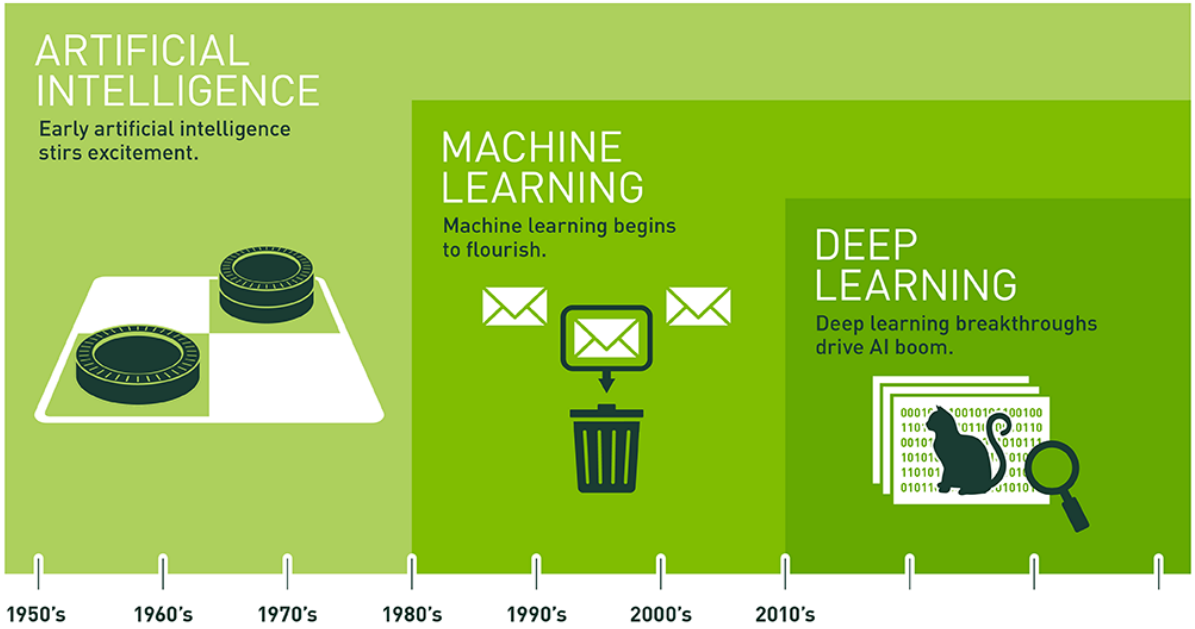
Often the weights, w_{ij}, are represented as a symmetric matrix $W = [w_{ij}]$, with zeros along the diagonal.

Boltzmann learning is similar to an error-correction learning rule, in that **an error signal is used to train the system in each iteration**. However, instead of a direct difference between the result value and the desired value, we take the difference between the probability distributions of the system.

Boltzmann Machine

- Boltzmann machines are used to solve two quite different computational problems.
- For a **search problem**, the weights on the connections are fixed and are used to represent a cost function. The stochastic dynamics of a Boltzmann machine then allow it to sample binary state vectors that have low values of the cost function.
- For a **learning problem**, the Boltzmann machine is shown a set of binary data vectors and it must learn to generate these vectors with high probability. To do this, it must find weights on the connections so that, relative to other possible binary vectors, the data vectors have low values of the cost function. To solve a learning problem, Boltzmann machines make many small updates to their weights, and each update requires them to solve many different search problems.

Deep Learning



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

Deep Learning

- Deep learning (also known as deep structured learning or hierarchical learning) is part of a broader family of machine learning methods based on learning data representations, as opposed to task-specific algorithms. Learning can be supervised, semi-supervised or unsupervised.
- Deep learning architectures such as deep neural networks, deep belief networks and recurrent neural networks have been applied to fields.
- Deep learning is a class of machine learning algorithms that:

1. use a cascade of multiple layers of nonlinear processing units for feature extraction and transformation. Each successive layer uses the output from the previous layer as input.
2. learn in supervised (e.g., classification) and/or unsupervised (e.g., pattern analysis) manners.
3. learn multiple levels of representations that correspond to different levels of abstraction; the levels form a hierarchy of concepts

Deep Learning

- Modern deep learning models are based on an artificial neural network (ANN).
- Applications of Deep Learning:

✓ Automatic speech recognition

✓ Image recognition

✓ NLP

✓ Image restoration & Military field

✓ Financial fraud detection.

THANK YOU

Any Queries ?