Machine Learning

Machine Learning

Definition of learning

- 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 T, as measured by P, improves with experience E.

Examples

- Handwriting recognition learning problem
  - Task T : Recognizing and classifying handwritten words within images
  - Performance P : Percent of words correctly classified
  - Training experience E : A dataset of handwritten words with given classifications
- A computer program which learns from experience is called a machine learning program or simply a learning program .

Classification of machine learning

Supervised learning

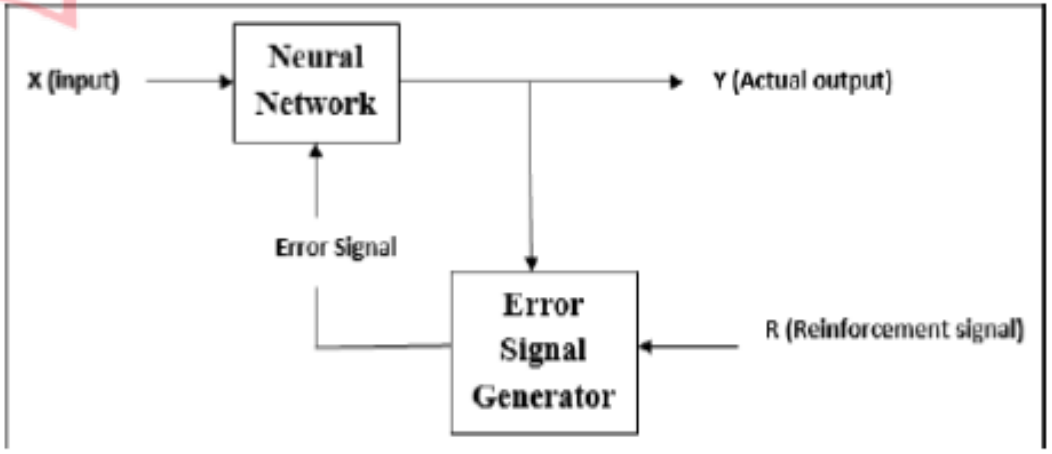
- As the name suggests, This learning process is dependent.
- During the training of ANN under supervised learning, all labelled input vector is presented to the network, which will give an output vector.
- This output vector is compared with the desired output vector. An error signal is generated, if there is a difference between the actual output and the desired output vector.
- Based on this error signal, the weights are adjusted until the actual output is matched with the desired output., which produces a mapping function, which can be used for mapping new examples i.e. future events.
- Classification algorithms like decision trees, Naive Bayes, k-nearest Neighbors and Support Vector Machines (SVM) are some examples of supervised learning.

Unsupervised learning

- As the name suggests, the learning process in unsupervised learning is independent.
- During the training of ANN, the input vectors of similar types are combined to form clusters.
- When a new input pattern is applied, then the neural network gives an output response indicating the class to which the input pattern belongs i.e. draws the inferences from datasets.
- There is no feedback from the environment as to what should be the desired output and if it is correct or incorrect.
- Hence, in this type of learning, the network itself must discover the patterns and features from the input data and the relation between the input data over the output.
- Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data.
- Clustering algorithms like K-means, mixture models, hierarchical clustering, anomaly detection, Generative Adversarial Networks etc. are some examples of unsupervised learning algorithms.

Reinforcement Learning (Assignment)

- As the name suggests, this type of learning is used to reinforce or strengthen the network over some critic information.
- The learning process is similar to supervised learning, however, we might have very little information.
- During the training of the network under reinforcement learning, the network receives some feedback from the environment. This makes it somewhat similar to supervised learning.
- However, the feedback obtained here is evaluative, not instructive, which means there is no teacher as in supervised learning. After receiving the feedback, the network performs adjustments of the weights to get better critic information in future.



Inductive Learning Algorithm?

- Inductive Learning Algorithm (ILA) is an iterative and inductive machine learning algorithm that is used for generating a set of classification rules, which produces rules of the form "IF-THEN", for a set of examples, producing rules at each iteration and appending to the set of rules.
- There are basically two methods for knowledge extraction firstly from domain experts and then with machine learning.
- For a very large amount of data, the domain experts are not very useful and reliable.
- So we move towards the machine learning approach for this work.
- To use machine learning One method is to replicate the expert's logic in the form of algorithms but this work is very tedious, time taking, and expensive.

Basic Requirements to Apply Inductive Learning Algorithm

- List the examples in the form of a table ‘T’ where each row corresponds to an example and each column contains an attribute value.
- Create a set of m training examples, each example composed of k attributes and a class attribute with n possible decisions.
- Create a rule set, R, having the initial value false.
- Initially, all rows in the table are unmarked.

Example

Example no.	Place type	weather	location	decision
1.	hilly	winter	kullu	Yes
2.	mountain	windy	Mumbai	No
3.	mountain	windy	Shimla	Yes
4.	beach	windy	Mumbai	No
5.	beach	warm	goa	Yes
6.	beach	windy	goa	No
7.	beach	warm	Shimla	Yes

Subset – 1

s.no	place type	weather	location	decision
1.	hilly	winter	kullu	Yes
2.	mountain	windy	Shimla	Yes
3.	beach	warm	goa	Yes
4.	beach	warm	Shimla	Yes

Subset – 2

s.no	place type	weather	location	decision
5.	mountain	windy	Mumbai	No
6.	beach	windy	Mumbai	No
7.	beach	windy	goa	No

- **At iteration 1** rows 3 & 4 column weather is selected and rows 3 & 4 are marked. the rule is added to R IF the weather is warm then a decision is yes.
- **At iteration 2** row 1 column place type is selected and row 1 is marked. the rule is added to R IF the place type is hilly then the decision is yes.
- **At iteration 3** row 2 column location is selected and row 2 is marked. the rule is added to R IF the location is Shimla then the decision is yes.
- **At iteration 4** row 5&6 column location is selected and row 5&6 are marked. the rule is added to R IF the location is Mumbai then a decision is no.
- **At iteration 5** row 7 column place type & the weather is selected and row 7 is marked. the rule is added to R IF the place type is beach AND the weather is windy then the decision is no.

Finally, we get the rule set:- Rule Set

- **Rule 1:** IF the weather is warm THEN the decision is yes.
- **Rule 2:** IF the place type is hilly THEN the decision is yes.
- **Rule 3:** IF the location is Shimla THEN the decision is yes.
- **Rule 4:** IF the location is Mumbai THEN the decision is no.
- **Rule 5:** IF the place type is beach AND the weather is windy THEN the decision is no.

Naïve Bayes Classifier Algorithm

- Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems.
- It is mainly used in *text classification* that includes a high-dimensional training dataset.
- This classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions.
- It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.
- Some popular examples of Naïve Bayes Algorithm are spam filtration, Sentimental analysis, and classifying articles.

Bayes' Theorem

- 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.
- The formula for Bayes' theorem is given as:

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

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.

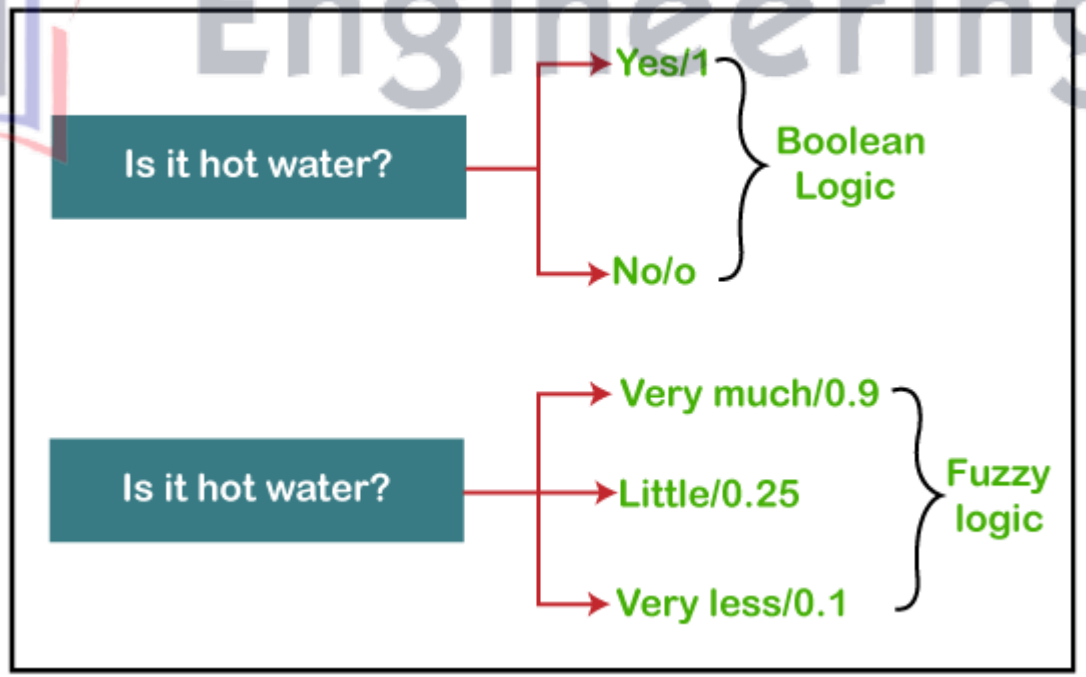
- **P(A) is Prior Probability**: Probability of hypothesis before observing the evidence.
- **P(B) is Marginal Probability**: Probability of Evidence.

Working of Naïve Bayes' Classifier

- Working of Naïve Bayes' Classifier can be understood with the help of the below example:
- Suppose we have a dataset of **weather conditions** and corresponding target variable "**Play**". So using this dataset we need to decide that whether we should play or not on a particular day according to the weather conditions. So to solve this problem, we need to follow the below steps:
- Convert the given dataset into frequency tables.
- Generate Likelihood table by finding the probabilities of given features.
- Now, use Bayes theorem to calculate the posterior probability.

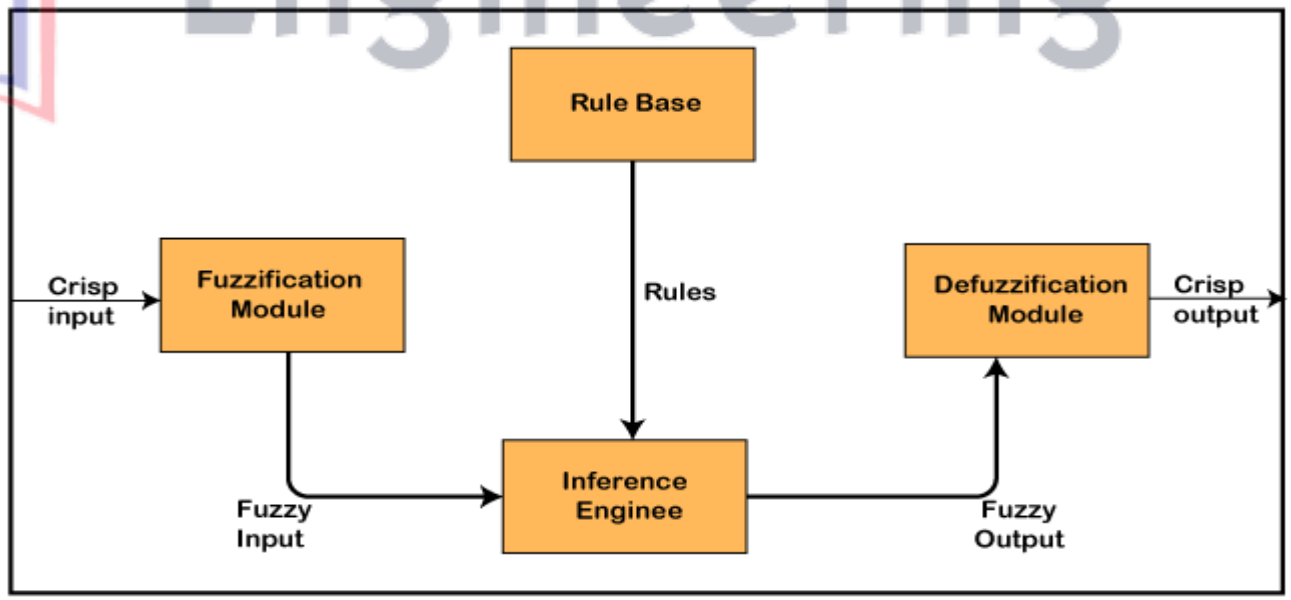
Fuzzy Logic

- The '**Fuzzy**' word means the things that are not clear or are vague.
- Sometimes, we cannot decide in real life that the given problem or statement is either true or false.
- At that time, this concept provides many values between the true and false and gives the flexibility to find the best solution to that problem.
- Example of Fuzzy Logic as comparing to Boolean Logic



Architecture of a Fuzzy Logic System

- In the architecture of the **Fuzzy Logic** system, each component plays an important role. The architecture consists of the different four components which are given below.
  - Rule Base
  - Fuzzification
  - Inference Engine
  - Defuzzification
- Following diagram shows the architecture or process of a Fuzzy Logic system:



Rule Base

- Rule Base is a component used for storing the set of rules and the If-Then conditions given by the experts are used for controlling the decision-making systems.
- There are so many updates that come in the Fuzzy theory recently, which offers effective methods for designing and tuning of fuzzy controllers.
- These updates or developments decreases the number of fuzzy set of rules.

Fuzzification

- **Fuzzification** is a module or component for transforming the system inputs, i.e., it converts the crisp number into fuzzy steps.
- The crisp numbers are those inputs which are measured by the sensors and then fuzzification passed them into the control systems for further processing.

- This component divides the input signals into following five states in any Fuzzy Logic system:
  - Large Positive (LP)
  - Medium Positive (MP)
  - Small (S)
  - Medium Negative (MN)
  - Large negative (LN)

Inference Engine

- This component is a main component in any Fuzzy Logic system (FLS), because all the information is processed in the Inference Engine.
- It allows users to find the matching degree between the current fuzzy input and the rules.
- After the matching degree, this system determines which rule is to be added according to the given input field.
- When all rules are fired, then they are combined for developing the control actions.

Defuzzification

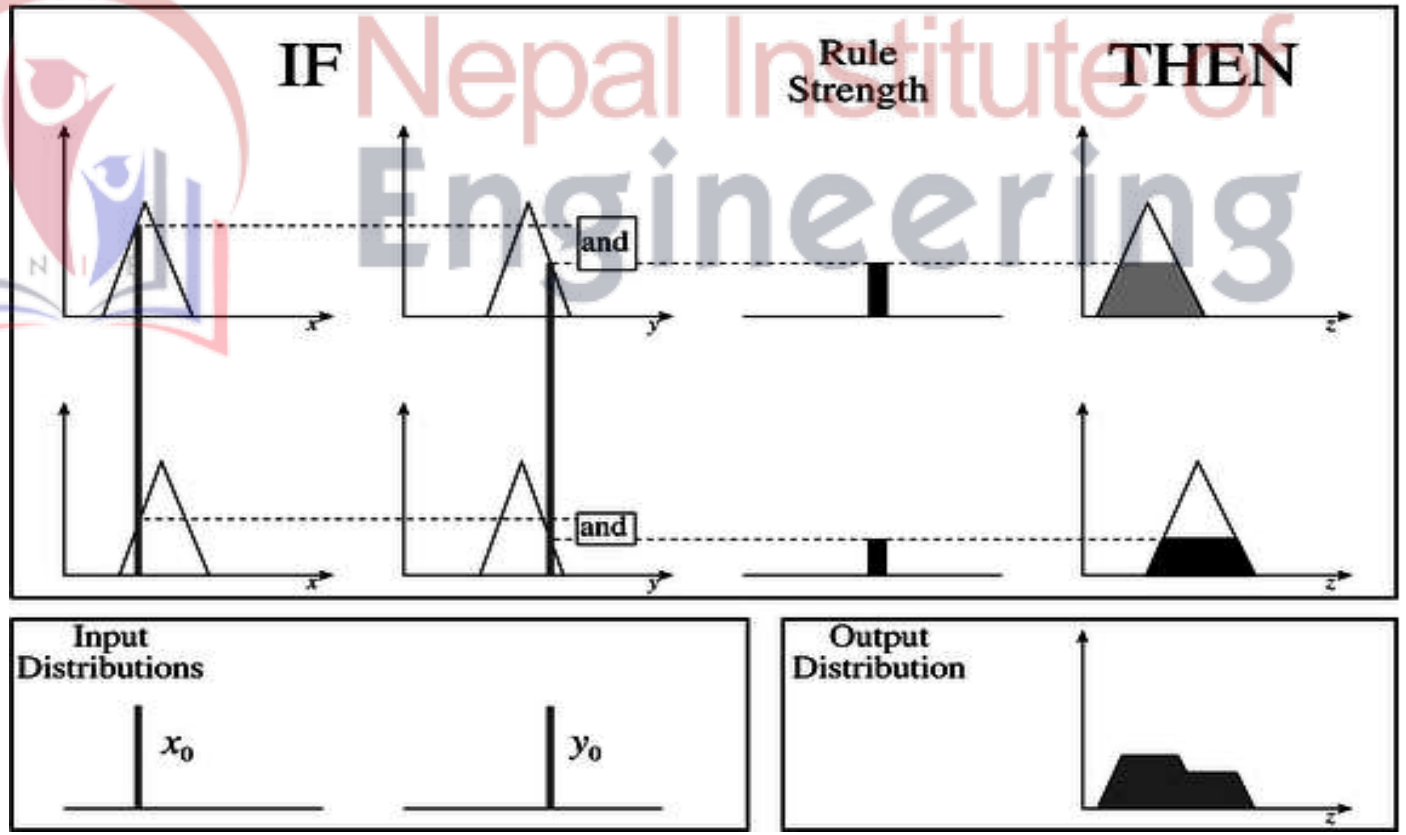
- **Defuzzification** is a module or component, which takes the fuzzy set inputs generated by the **Inference Engine**, and then transforms them into a crisp value.
- It is the last step in the process of a fuzzy logic system.
- The crisp value is a type of value which is acceptable by the user.
- Various techniques are present to do this, but the user has to select the best one for reducing the errors.

Methods of FIS (Fuzzy Inference System)

- Following are the two important methods of FIS, having different consequent of fuzzy rules −
  - Mamdani Fuzzy Inference System
  - Takagi-Sugeno Fuzzy Model (TS Method)

Mamdani Fuzzy Inference System

- This system was proposed in 1975 by Ebhasim Mamdani. Basically, it was anticipated to control a steam engine and boiler combination by synthesizing a set of fuzzy rules obtained from people working on the system.
- Block diagram of Mamdani Fuzzy Inference System



Takagi-Sugeno Fuzzy Model (TS Method)

- This model was proposed by Takagi, Sugeno and Kang in 1985. Format of this rule is given as −
- *IF x is A and y is B THEN Z = f(x,y)*
- Here, *AB* are fuzzy sets in antecedents and *z = f(x,y)* is a crisp function in the consequent.

Genetic Algorithm

- A genetic algorithm is an adaptive heuristic search algorithm inspired by "Darwin's theory of evolution in Nature."

How Genetic Algorithm Work?

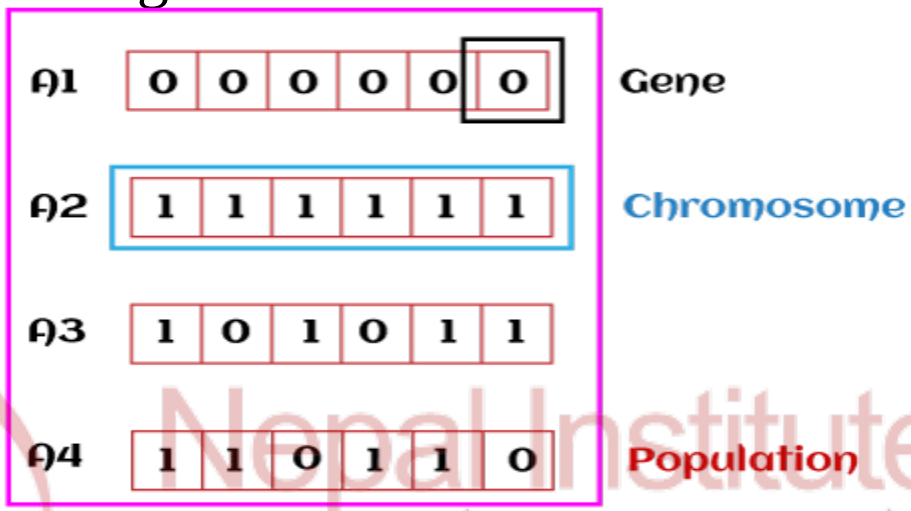
- The genetic algorithm works on the evolutionary generational cycle to generate high-quality solutions.
- These algorithms use different operations that either enhance or replace the population to give an improved fit solution.

- It basically involves five phases to solve the complex optimization problems, which are given as below:
  - **Initialization**
  - **Fitness Assignment**
  - **Selection**
  - **Reproduction**
  - **Termination**

Initialization

- The process of a genetic algorithm starts by generating the set of individuals, which is called population.
- Here each individual is the solution for the given problem.
- **An individual contains or is characterized by a set of parameters called Genes.**
- Genes are combined into a string and generate chromosomes, which is the solution to the problem.

- One of the most popular techniques for initialization is the use of random binary strings.



Fitness Assignment

- Fitness function is used to determine how fit an individual is?
- It means the ability of an individual to compete with other individuals.
- In every iteration, individuals are evaluated based on their fitness function.
- The fitness function provides a fitness score to each individual.

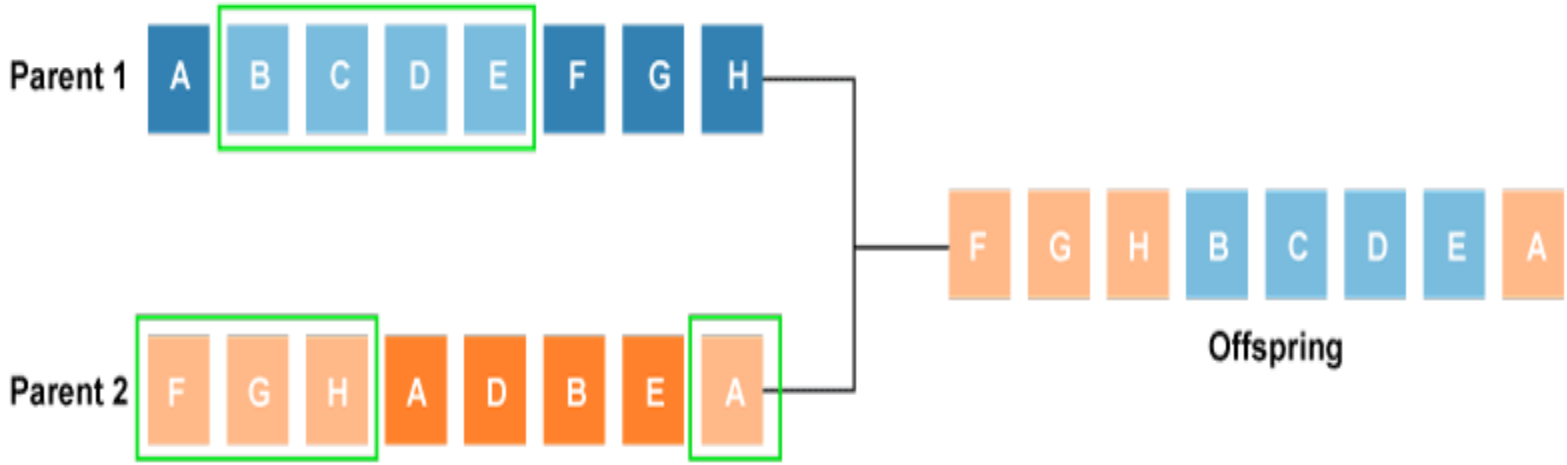
- This score further determines the probability of being selected for reproduction.
- **The high the fitness score, the more chances of getting selected for reproduction.**

Selection

- The selection phase involves the selection of individuals for the reproduction of offspring.
- All the selected individuals are then arranged in a pair of two to increase reproduction.
- Then these individuals transfer their genes to the next generation.
- There are three types of Selection methods available, which are
  - Roulette wheel selection
  - Tournament selection
  - Rank-based selection

Reproduction

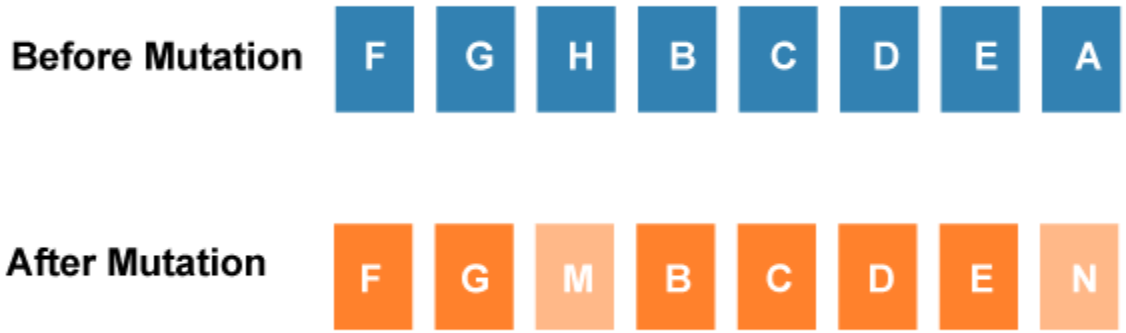
- After the selection process, the creation of a child occurs in the reproduction step.
- In this step, the genetic algorithm uses two variation operators that are applied to the parent population.
- The two operators involved in the reproduction phase are given below:
  - **Crossover**
    - The crossover plays a most significant role in the reproduction phase of the genetic algorithm.
    - In this process, **a crossover point is selected at random within the genes.**
    - Then **the crossover operator swaps genetic information of two parents from the current generation to produce a new individual representing the offspring.**



- The genes of parents are exchanged among themselves until the crossover point is met.
- These newly generated offspring are added to the population. This process is also called or crossover.
- Types of crossover styles available:
  - One point crossover
  - Two-point crossover
  - Livery crossover
  - Inheritable Algorithms crossover

Mutation

- The mutation operator inserts random genes in the offspring (new child) to maintain the diversity in the population.
- It can be done by flipping some bits in the chromosomes.
- Mutation helps in solving the issue of premature convergence and enhances diversification.
- The below image shows the mutation process: Types of mutation styles available,
  - **Flip bit mutation**
  - **Gaussian mutation**
  - **Exchange/Swap mutation**



Termination

- After the reproduction phase, a stopping criterion is applied as a base for termination.
- The algorithm terminates after the threshold fitness solution is reached.
- It will identify the final solution as the best solution in the population.

- What is Machine learning?
  a) The selective acquisition of knowledge through the use of computer programs
  b) The selective acquisition of knowledge through the use of manual programs
  c) The autonomous acquisition of knowledge through the use of computer programs
  d) The autonomous acquisition of knowledge through the use of manual programs
- Which of the following is not a supervised machine learning algorithm?
  a) K-means
  b) Naïve Bayes
  c) SVM for classification problems
  d) Decision tree

- What is the key difference between supervised and unsupervised learning?
  a) Supervised learning requires labeled data, while unsupervised learning does not.
  b) Supervised learning predicts labels, while unsupervised learning discovers patterns.
  c) Supervised learning is used for classification, while unsupervised learning is used for regression.
  d) Supervised learning is always more accurate than unsupervised learning.
- Which one of the following models is a generative model used in machine learning?
  a) Support vector machines
  b) Naïve Bayes
  c) Logistic Regression
  d) Linear Regression

- Which combines inductive methods with the power of first-order representations?
  a) Inductive programming
  b) Logic programming
  c) Inductive logic programming
  d) Lisp programming
- How many reasons are available for the popularity of ILP?
  a) 1
  b) 2
  c) 3
  d) 4

Knowledge, hypothesis, and algorithm

- Which is an appropriate language for describing the relationships?
  a) First-order logic
  b) Propositional logic
  c) ILP
  d) None of the mentioned

- Which produces hypotheses that are easy to read for humans?
  a) ILP
  b) Artificial intelligence
  c) Propositional logic
  d) First-order logic
- What is the form of Fuzzy logic?
  a) Two-valued logic
  b) Crisp set logic
  c) Many-valued logic
  d) Binary set logic
- The truth values of traditional set theory is _____ and that of fuzzy set is _____
  a) Either 0 or 1, between 0 & 1
  b) Between 0 & 1, either 0 or 1
  c) Between 0 & 1, between 0 & 1
  d) Either 0 or 1, either 0 or 1

- The room temperature is hot. Here the hot (use of linguistic variable is used) can be represented by _____
  a) Fuzzy Set
  b) Crisp Set
  c) Fuzzy & Crisp Set
  d) None of the mentioned
- The values of the set membership is represented by _____
  a) Discrete Set
  b) Degree of truth
  c) Probabilities
  d) Both Degree of truth & Probabilities
- Fuzzy logic is usually represented as _____
  a) IF-THEN-ELSE rules
  b) IF-THEN rules
  c) Both IF-THEN-ELSE rules & IF-THEN rules
  d) None of the mentioned

- What will happen when a chain-termination mutation is found in the S gene?
  - Cell lysis gets blocked.
  - The growth of cells containing low levels of packaging proteins is not allowed.
  - The lysis of cells cannot be carried artificially
  - Packaging cannot be carried out efficiently
- What protects the modified DNA after undergoing replication?
  - Conservative mode of replication
  - Semi-conservative mode of replication
  - Replication has nothing to do with replication
  - DNA gets modified after replication
- What is the name of the phenomenon in which the specificity of the enzyme gets affected by the concentration of the used buffer?
  - specificity elevation
  - diamond activity
  - star activity
  - concentration gradient effects

Genetic algorithm involves

- Mutation
- Crossover
- Selection
- All