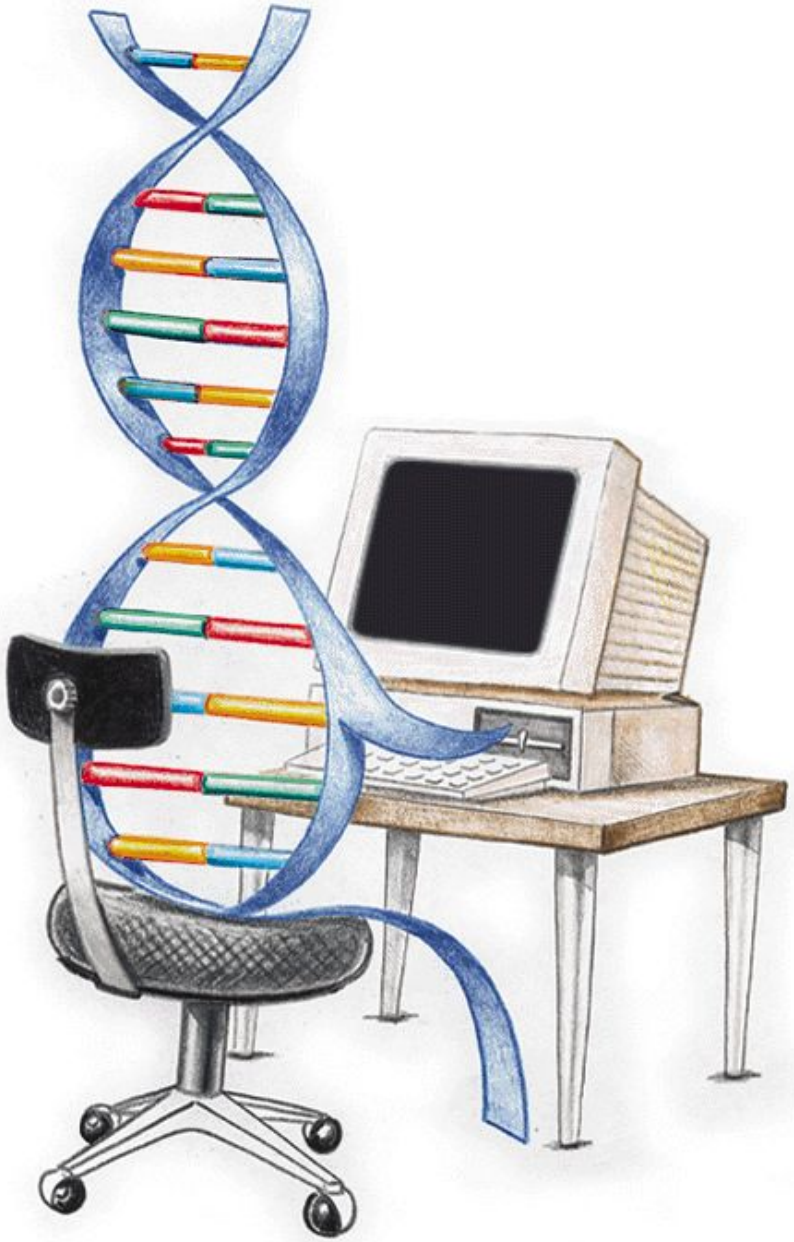


Genetic Algorithms in Machine Learning

Machine Learning – is the study of **computer algorithms** that improve automatically through experience. It is seen as a subset of **artificial intelligence**

Supervised ML

- Task driven
- Predict next value
- **Classification Algorithm** (eg: Fraud detection, Email Spam detection, Image Classification, Diagnostics)
- **Regression Algorithm** (eg: Risk Assessment , Score prediction)

Unsupervised ML

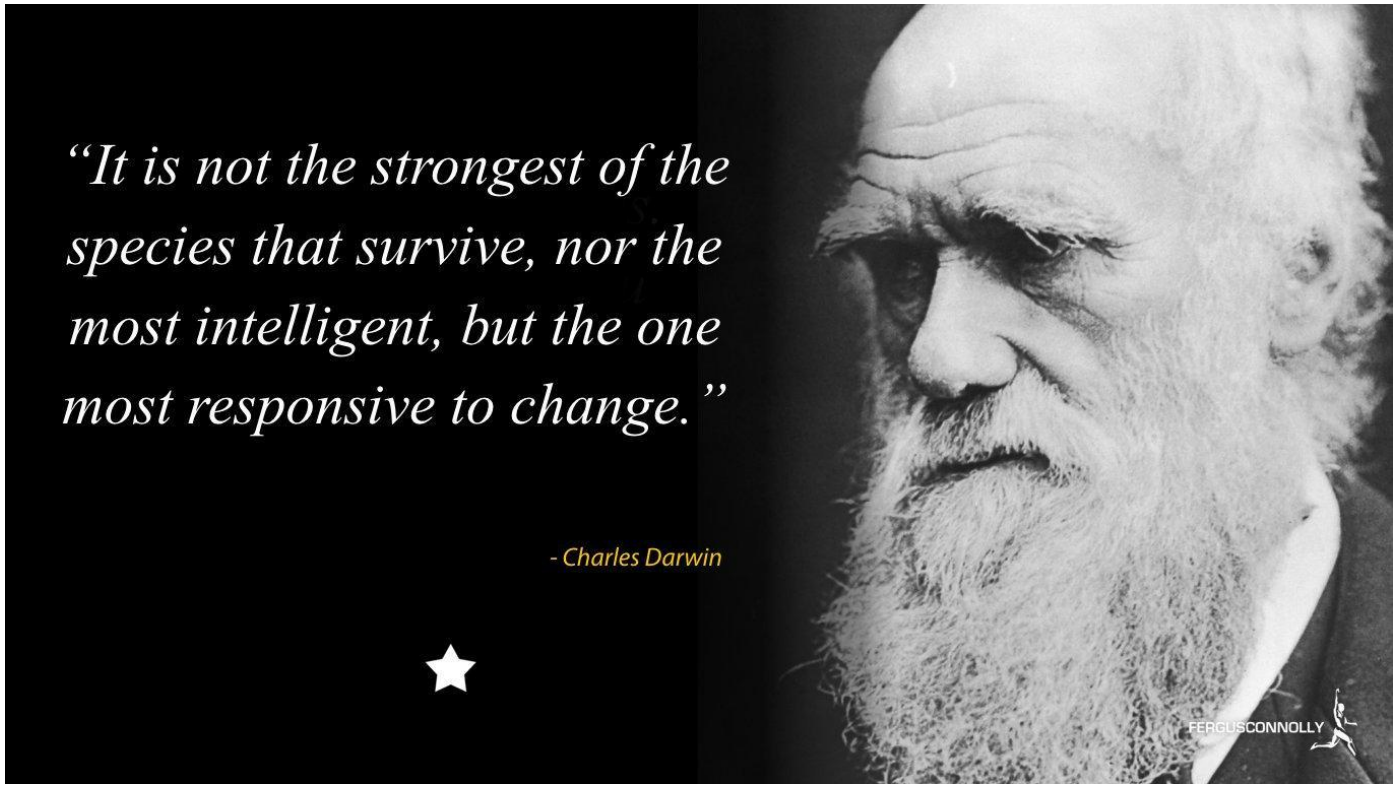
- Data driven
- Identify Clusters
- **Clustering Algorithm** (eg: Biology, City planning, Targeted Marketing)
- **Dimensionality reduction** (eg: Text Mining, Face/Image recognition, Big data visualization, compression techniques)

Reinforcement ML

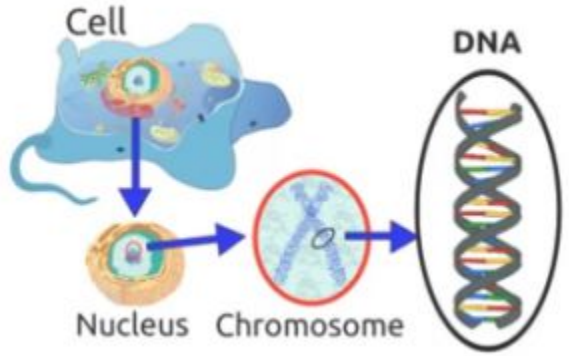
- Learn from mistakes and react to the environment
- eg: Gaming, Financial sectors, Inventory management, Manufacturing, Robotic automation/navigation, Real time AI-like Alexa, driverless CARS

In machine learning the computer learns

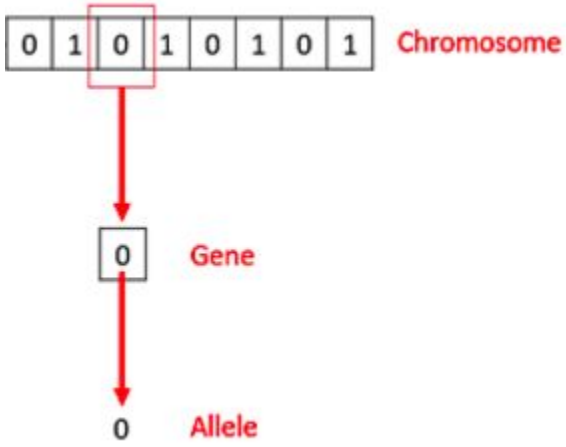
1. Intuition behind Genetic Algorithms



2. Biological Inspiration



- ***Cells*** *are the basic* ***building block*** *of all living things.*
- Therefore in each cell, there is the same set of chromosomes.
  - **Chromosome** are basically the strings of DNA.
- Traditionally, these chromosomes are represented in binary as strings of 0's and 1's.
  - A chromosome consists of **genes**, commonly referred as blocks
    of DNA, where each gene encodes a specific trait,
    for example hair color or eye color.



3. Simple example of GA

Consider a hypothetical situation where, you are head of a country, and in order to keep your city safe from bad things, you implement a policy like this.

You select all the good people, and ask them to extend their generation by having their children. This repeats for a few generations.

You will notice that now you have an entire population of good people.

GA mechanism

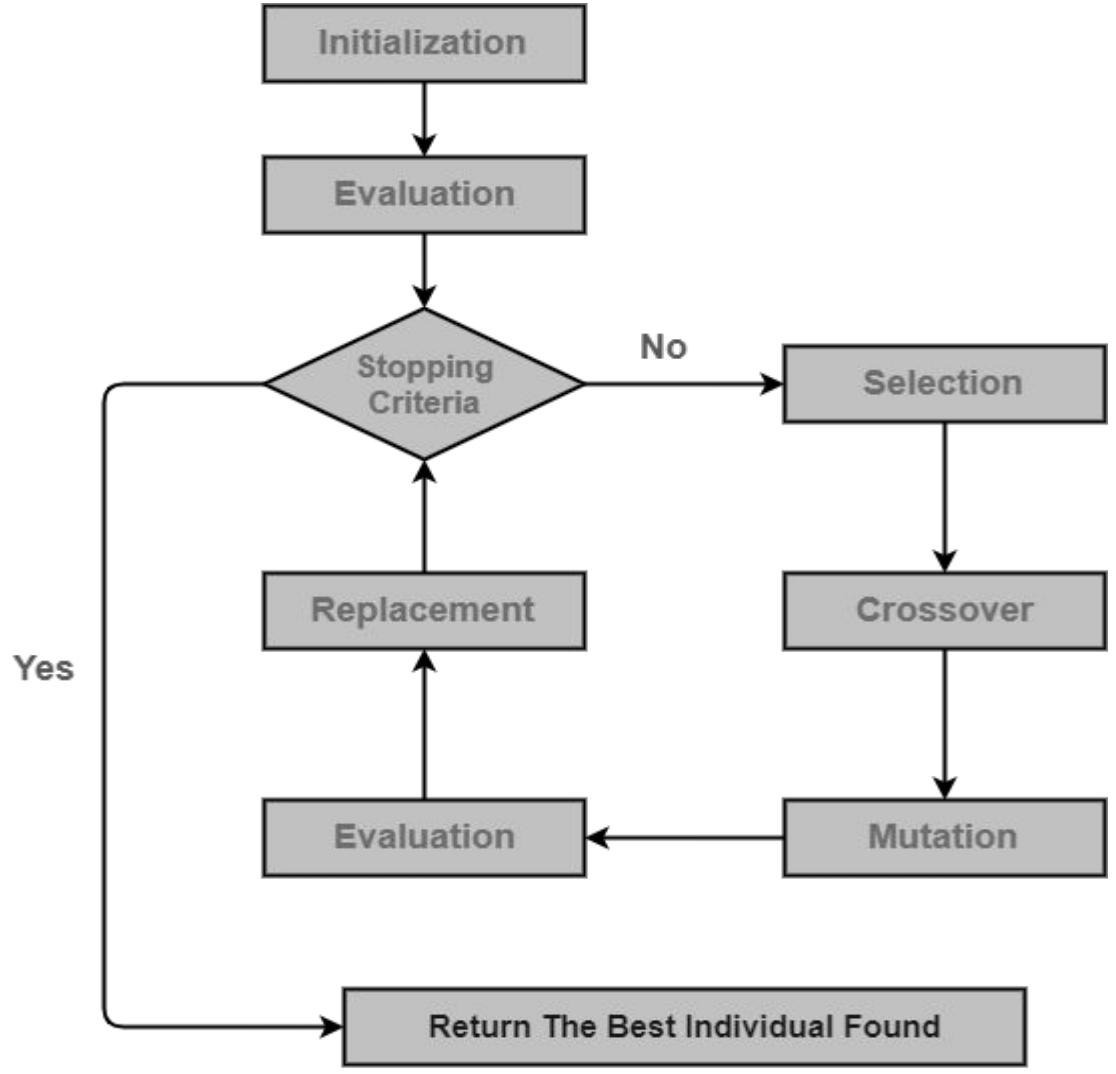
- Firstly, we defined our initial population as our countrymen.
- We defined a function to classify whether is a **person is good or bad.**
- Then we **selected** good people for mating to produce their off-springs.
- And finally, these off-springs **replace** the bad people from the population and this process repeats.
- This is how genetic algorithm actually works, which basically tries to **mimic the human evolution** to some extent.
- it is an **optimization technique**, which tries to find out such values of input so that we get the best output values or results.
- The working of a genetic algorithm is also derived from biology

GENETIC ALGORITHM [GA]

- A genetic algorithm is a **search heuristic** that mimics the process of natural evolution used to generate useful solutions to optimization and search problems.
- Genetic algorithms are a subset of **evolutionary algorithms** which solves optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover
- **Crossover** and **mutation** are two basic operators of GA.
- Performance of GA very depend on them.
- Type and implementation of operators depends on encoding and also on a problem.

Five phases in a genetic algorithm.

- Initial population
- Fitness function
- Selection
- Crossover
- Mutation



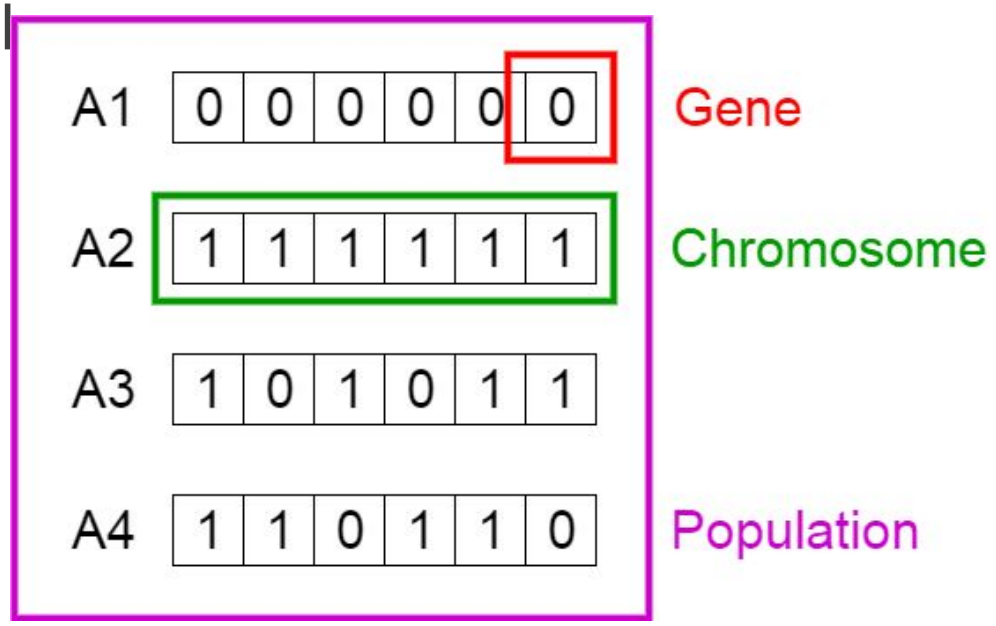
How Genetic Algorithm Works

Phase 1: Initial Population

The process begins with a **set of individuals** which is called a **Population**. Each individual is a solution to the problem you want to solve.

An individual is characterized by a **set of parameters (variables)** known as **Genes**. Genes are joined into a string to form a **Chromosome** (solution).

In a genetic algorithm, the set of genes of an individual is represented using a string, in terms of an alphabet. Usually, binary values are used (string of 1s and 0s). We say that we encode the genes in a ch



Phase 2:Fitness function

The **fitness function** determines how fit an individual is (the ability of an individual to compete with other individuals). It gives a **fitness score** to each individual. The probability that an individual will be selected for reproduction is based on its fitness score.

Phase 3: Selection

The idea of **selection** phase is to select the fittest individuals and let them pass their genes to the next generation.

Two pairs of individuals (**parents**) are selected based on their fitness scores. Individuals with high fitness have more chance to be selected for reproduction.

Selection

Chromosomes are selected from the population to be parents to crossover.

According to Darwin's evolution theory the best ones should survive and create new offspring.

There are many methods how to select the best chromosomes, for example roulette wheel selection

Roulette Wheel Selection : Parents are selected according to their **fitness.** The better the chromosomes are, the more chances to be selected they have. Imagine a roulette wheel where are placed all chromosomes in the population, every chromosome has its place big accordingly to its fitness function

Crossover is the most significant phase in a genetic algorithm. For each pair of parents to be mated, a **crossover point** is chosen at random from within the genes.

Types of Crossover

1) Single point crossover - In this case one crossover point is selected, binary string from beginning of chromosome to the crossover point is copied from one parent, and the rest is copied from the second parent 11001011+11011111 = 11001111

2) Two point crossover - Here two crossover points are selected, binary string from beginning of chromosome to the first crossover point is copied from one parent, the part from the first to the second crossover point is copied from the second parent and the rest is copied from the first parent 11001011 + 11011111 = 11011111

3) Uniform crossover - In this method bits are randomly copied from the first or from the second parent 11001011 + 11011101 = 11011111

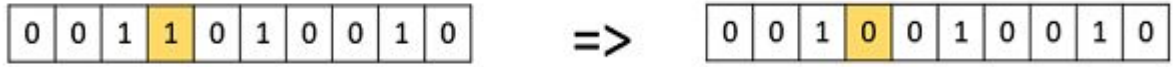
Phase 5: Mutation

- Mutation is a genetic operator used to **maintain genetic diversity** from one generation of a population to the next.
- It is similar to biological mutation.
- In simple terms, mutation may be defined as a **small random tweak** in the chromosome, to get a new solution.
- An example of a mutation operator involves a probability that an arbitrary bit in a genetic sequence will be changed from its original state.
- A common method of implementing the mutation operator involves generating a random variable for each bit in a sequence.
- The purpose of mutation in GAs is **preserving and introducing diversity.**
- Mutation should allow the algorithm to **avoid local minima** by preventing the population of chromosomes from becoming too similar to each other.

Types of Mutation

Bit Flip Mutation

In this bit flip mutation, we select one or more random bits and flip them. This is used for binary encoded GAs.

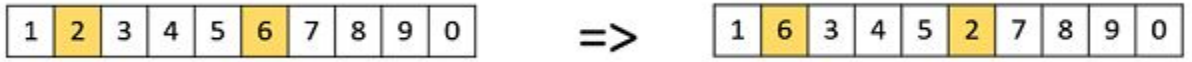


Random Resetting

Random Resetting is an extension of the bit flip for the integer representation. In this, a random value from the set of permissible values is assigned to a randomly chosen gene.

Swap Mutation

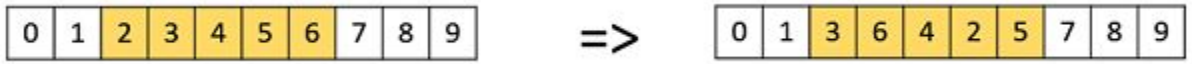
In swap mutation, we select two positions on the chromosome at random, and interchange the values. This is common in permutation based encodings.



Types of Mutation [contd..]

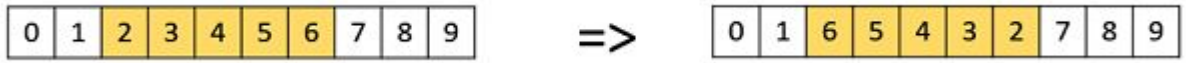
Scramble Mutation

Scramble mutation is also popular with permutation representations. In this, from the entire chromosome, a subset of genes is chosen and their values are scrambled or shuffled randomly.



Inversion Mutation

In inversion mutation, we select a subset of genes like in scramble mutation, but instead of shuffling the subset, we merely invert the entire string in the subset.



Applications of GA

Optimization − to maximize or minimize a given objective function value under a given set of constraints.

Economics − cobweb model, game theory equilibrium resolution, asset pricing, etc.

Neural Networks − to train neural networks, particularly recurrent neural networks.

Parallelization

Image Processing − DIP tasks as well like dense pixel matching.

Vehicle routing problems − fleet management system.

Scheduling applications − the time tabling problem.

Machine Learning − **genetics based machine learning (GBML)** is a niche area in machine learning.

Robot Trajectory Generation − to plan the path which a robot arm takes by moving from one point to another.

Parametric Design of Aircraft − GAs have been used to design aircrafts by varying the parameters and evolving better solutions.

DNA Analysis − to determine the structure of DNA using spectrometric data about the sample.

Multimodal Optimization − GAs are obviously very good approaches for multimodal optimization in which we have to find multiple optimum solutions.

Traveling salesman problem and its applications

Pseudocode for GA

START

Generate the initial population

Compute fitness

REPEAT

    Selection

    Crossover

    Mutation

    Compute fitness

UNTIL population has converged

STOP

Thank you