Pune Institute of Computer Technology,

Dhankadi,Pune

DEPARTMENT OF COMPUTER ENGINEERING

A

Project Report

On

“Performing Genetic Algorithm on Iris Dataset”

Submitted by

Aakash Patil (41247)

Saket Gupta (41256)

Sharlene Wadhwa (41260)

In partial fulfillment of

BACHELOR OF ENGINEERING

In

Computer Engineering

Savitribai Phule Pune University, Pune

INTRODUCTION

Our task is to determine the optimal categories on Iris Dataset using Genetic Algorithm. We accept various parameters from the user. And display sets of genetically generated results. These results may or may not be optimal.

OBJECTIVE

To Apply Genetic Algorithm for Iris Dataset.

H/W AND S/W REQUIREMENTS

Hardware Requirements: PIV, 2GB RAM, 500 GB HDD, Lenovo A13-4089Model

Software Requirements : Anaconda with Python 3.7

THEORY CONCEPTS

These algorithms can be implemented to find a solution to the optimization problems of various types. One such problem is determining categories based on various parameters on Iris Dataset.

Approach: In the following implementation, categories are taken as genes, parameters to be followed is the chromosome. After performing genetic algorithm, we obtain list of child solutions.

Operators Used:

1. Selection – Tournament selection without replacement
2. Crossover – Order crossover with window size 3
3. Mutation – Swap mutation

Algorithm:

1. Initialize the population randomly.
2. Determine the fitness of the chromosome.
3. Until done repeat:
    1. Select parents.
    2. Perform crossover and mutation.
    3. Calculate the fitness of the new population.
    4. Append it to the gene pool.

Advantages of GA's

• Does not require any derivative information (which may not be available for many real- world problems).

• Is faster and more efficient as compared to the traditional methods. • Has very good parallel capabilities.

• Optimizes both continuous and discrete functions and also multi-objective problems. • Provides a list of "good" solutions and not just a single solution.

• Always gets an answer to the problem, which gets better over the time.

• Useful when the search space is very large and there are a large number of parameters involved.

Limitations of GA's

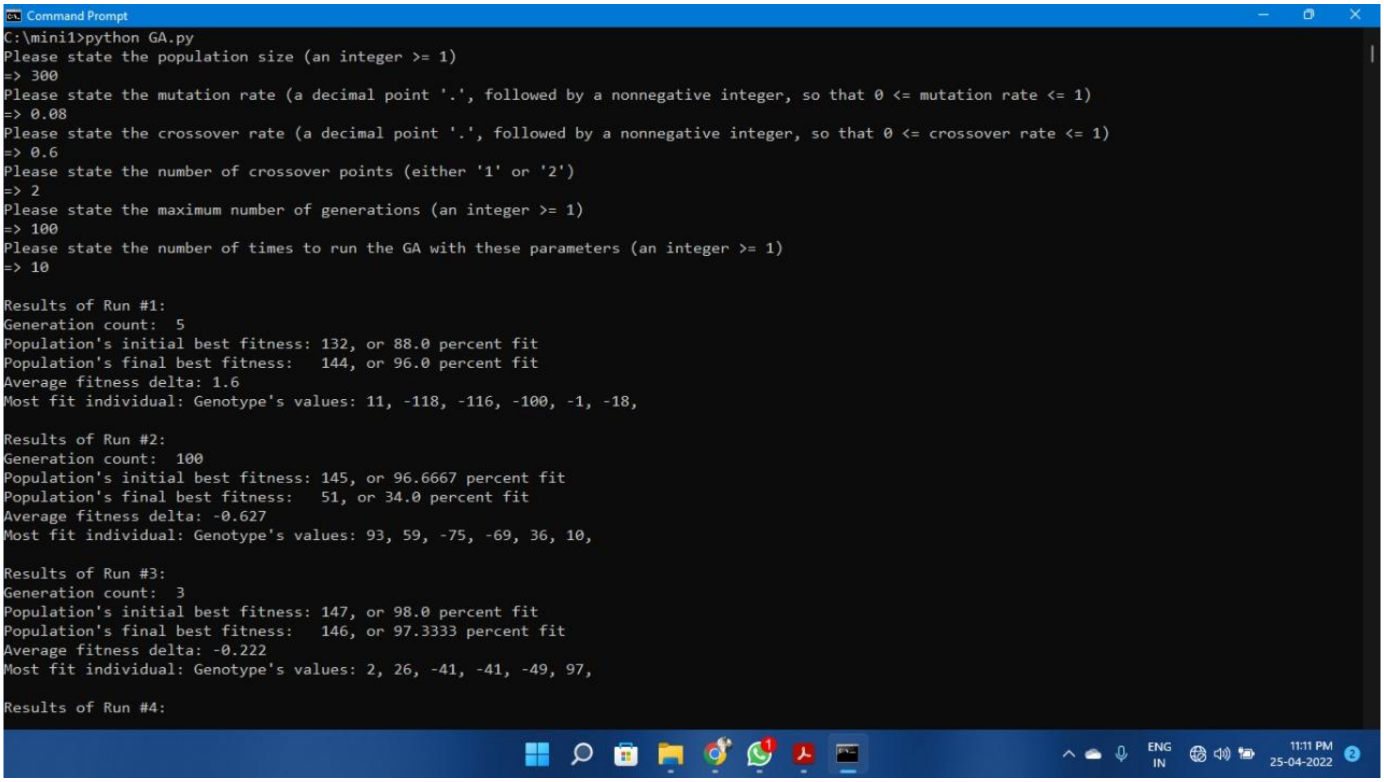
• GAs are not suited for all problems, especially problems which are simple and for which derivative information is available.

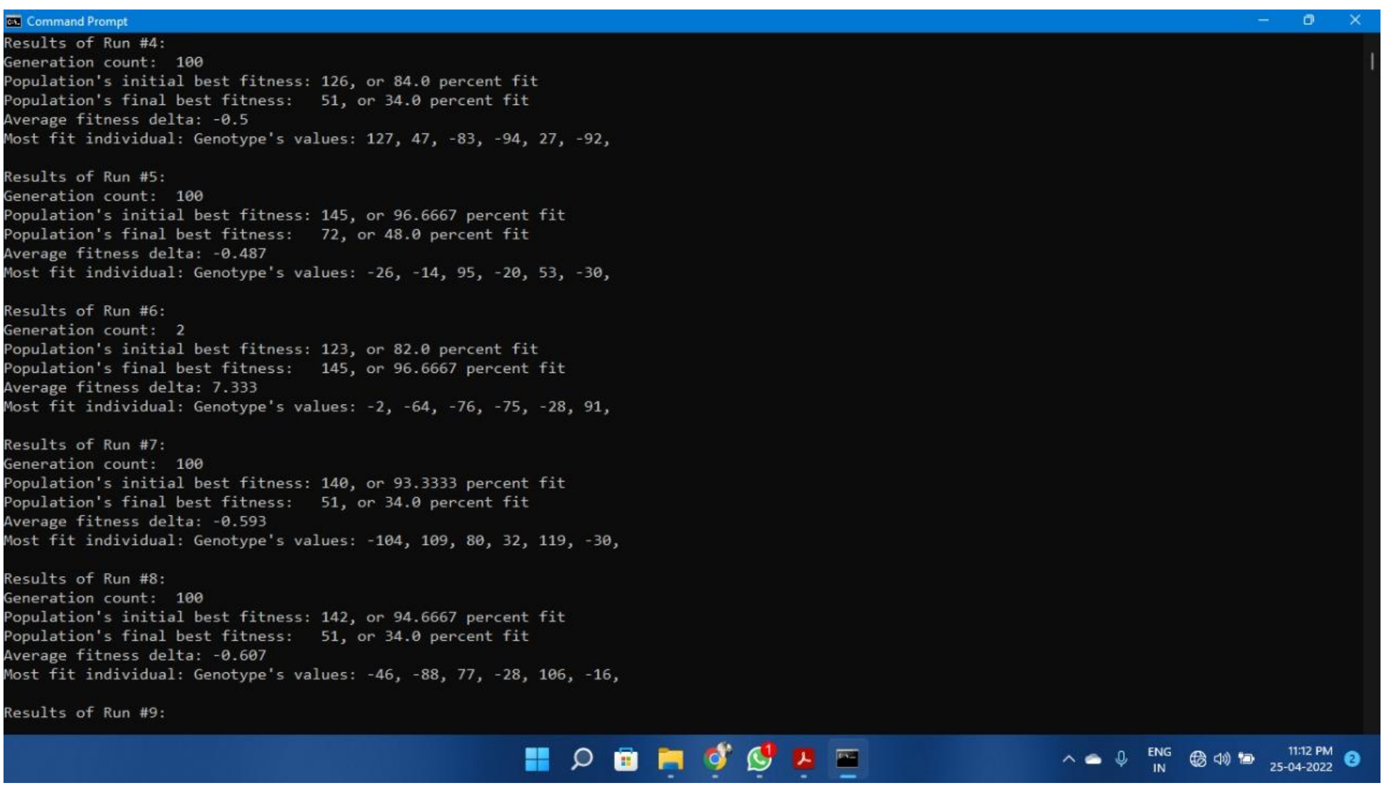
• Fitness value is calculated repeatedly which might be computationally expensive for some problems.

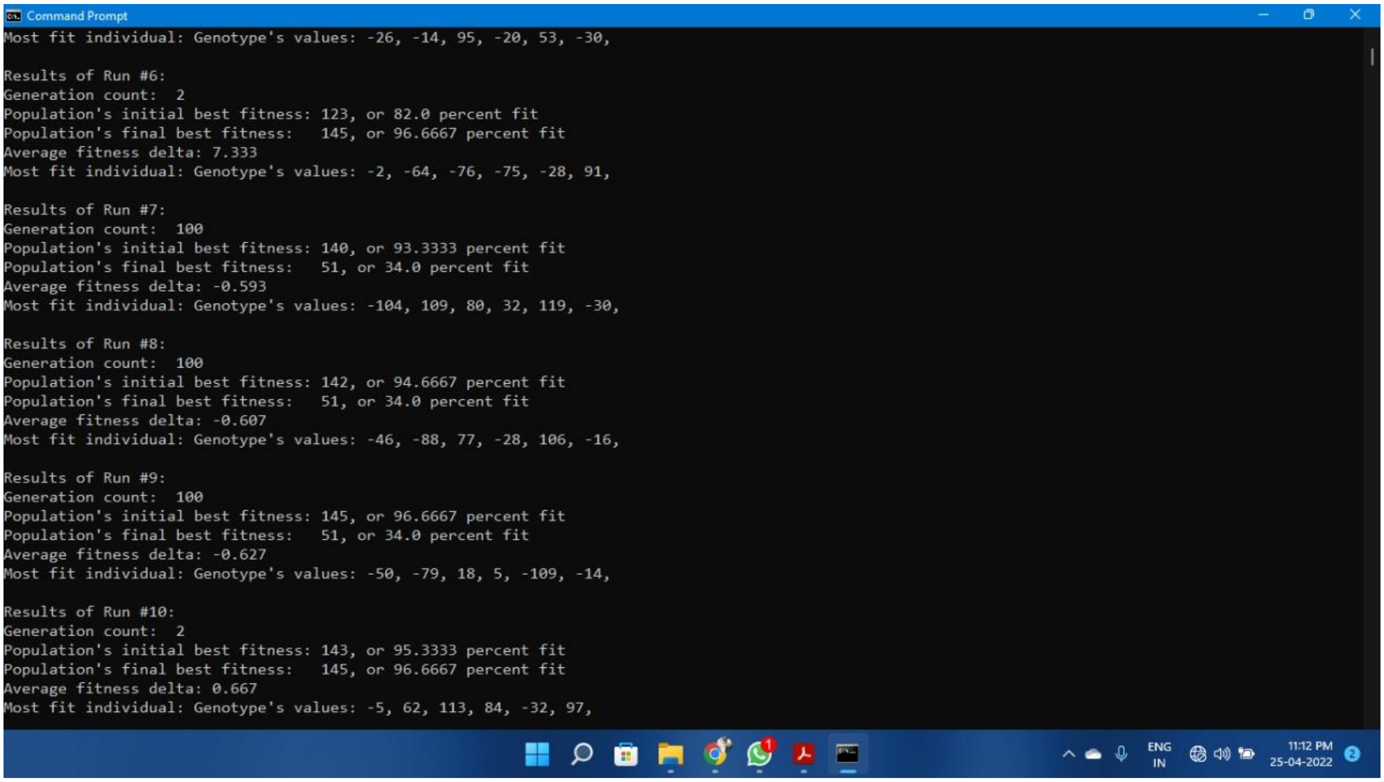
• Being stochastic, there are no guarantees on the optimality or the quality of the solution.

• If not implemented properly, the GA may not converge to the optimal solution.

SCREENSHOTS







```
Command Prompt
Most fit individual: Genotype's values: -26, -14, 95, -20, 53, -30,

Results of Run #6:
Generation count:  2
Population's initial best fitness: 123, or 82.0 percent fit
Population's final best fitness:   145, or 96.6667 percent fit
Average fitness delta: 7.333
Most fit individual: Genotype's values: -2, -64, -76, -75, -28, 91,

Results of Run #7:
Generation count:  100
Population's initial best fitness: 140, or 93.3333 percent fit
Population's final best fitness:   51, or 34.0 percent fit
Average fitness delta: -0.593
Most fit individual: Genotype's values: -104, 109, 80, 32, 119, -30,

Results of Run #8:
Generation count:  100
Population's initial best fitness: 142, or 94.6667 percent fit
Population's final best fitness:   51, or 34.0 percent fit
Average fitness delta: -0.607
Most fit individual: Genotype's values: -46, -88, 77, -28, 106, -16,

Results of Run #9:
Generation count:  100
Population's initial best fitness: 145, or 96.6667 percent fit
Population's final best fitness:   51, or 34.0 percent fit
Average fitness delta: -0.627
Most fit individual: Genotype's values: -50, -79, 18, 5, -109, -14,

Results of Run #10:
Generation count:  2
Population's initial best fitness: 143, or 95.3333 percent fit
Population's final best fitness:   145, or 96.6667 percent fit
Average fitness delta: 0.667
Most fit individual: Genotype's values: -5, 62, 113, 84, -32, 97,
```


CONCLUSION

Hence ,Successfully studied and practically implemented Genetic Algorithm on Iris Datase