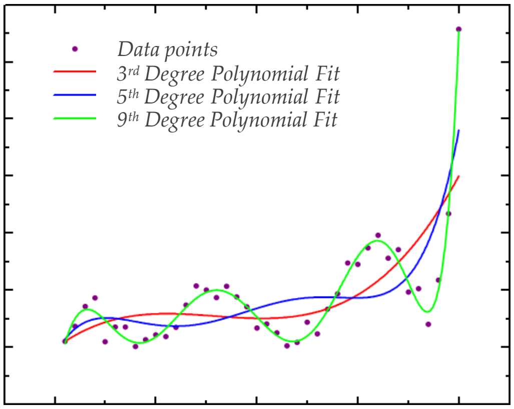
Cairo University
Faculty of Computers & Artificial Intelligence
Department of Computer Science
Soft Computing Course - 2023

Assignment 2 – Smooth Curve Fitting Problem

About the problem:

Curve fitting is the process of constructing a curve, or mathematical function (polynomial equation) that has the best fit to a series of data points, possibly subject to constraints. In smooth curve fitting, the function is constructed to approximately fit the data. Given a set of points, we would like to fit a curve to them using a polynomial equation.



What you are required to do:

Write a genetic algorithm to find the best coefficients that would make the distance between the polynomial function and the points minimum. (A solved example is provided in lab Floating point genetic algorithms)

What the input looks like:

You'll be given an input file with the following format:

- *First line:* Number of datasets (must be at least 1)
  For each dataset:
- Number of data points and polynomial degree separated by space
  For each point:
- x-coordinate and y-coordinate separated by space

Example:

```
1

4 2
1 5
2 8
3 13
4 20
```

Note: In the example above, we have 1 dataset containing 4 points and we want to fit a 2nd degree polynomial **(a_0, a_1, a_2)**.

Important remarks:

1. The first thing you need to do is to think about how you will encode the solution in your chromosome and what the objective function will be.
2. You should use a **floating-point representation** of the chromosome.
3. You can try different population sizes to see how this will affect your results. The maximum number of generations is up to you.
4. Initialize the genes such that their values are in **the range [-10,10].**
5. Implement **tournament selection**, **2-point crossover, non-uniform mutation, and elitist replacement**.
6. You must **read the input from the given file and write the output to a file.** The output should consist of the dataset index, the coefficients of the polynomial function, and its mean square error.

Assignment submission notes:

- The **maximum** number of students in a team is **3** and the **minimum** is **2 from the same lab**.
- You could use any programming language.
- The deadline will be on Wednesday **15/11 at 11:59 pm**, and no late submission is allowed.
- Please submit one compressed folder. The folder name should follow this structure: ID1_ID2_ID3_MAJOR _GROUP
- Cheating students will take negative grades and no excuses will be accepted. If you have any problems during the submission, contact your TA but don't, under any circumstances, give your code to or take the code from your friends.

Grading Criteria: (5 marks)

Representation and initialization	0.5
Fitness function	1
Selection	0.75
Crossover	0.75
Mutation	0.5
Replacement	0.5
Output	1

Good luck