Practice: Feature Selection with Evolutionary Algorithms

Feature selection involves choosing a subset of features in a Machine Learning problem. The main idea is to keep or increase the model performance while the number of features decreases.

Individual's representation

Assuming we have a problem with D features, feature selection consists of selecting only k of them, where k<D. The individual's representation could be a binary vector with D elements, whose entries represent if the i-th feature is or not selected. The search space is equal to 2^D.

Feature 1	Feature 2	…	…	…	…	…	…	Feature D
0	1	1	1	0	0	0	0	1

Individual's fitness

We are going to use two optimization criteria:

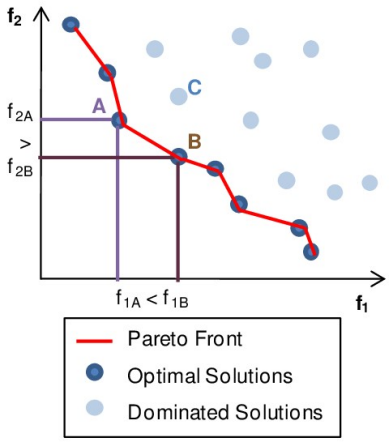
- F_1 = Minimize the error of the learning model
- F_2 = Minimize the number of features

In this sense, we will have dominated and non-dominated solutions. A dominated solution is a solution that exists at least another solution with better values in the two objectives. A non-dominated solution is one, in the Pareto font, that, compared to others, there not exist another solution better in the two objective functions.

For this feature selection problem, we will define that solution A dominates solution B if:

- E(A) < E(B)
- F(A) <= F(B)

where E(.) corresponds to the error of the learning model and F(.) corresponds to the number of features.



Parent selection

In this case, we will use a binary tournament. If one individual dominates the other, we will select the non-dominant. Else, we will select the one with the smaller value in the error of the learning model.

Evaluation	
Element	Points
Step 1: Create the initial population	5
Step 2: Create a function for calculating the individual's fitness and plot individuals in a 2d map.	15
Step 3: Implement the parent selection function	10
Step 4: Implement the crossover function	10
Step 5: Implement the mutation function	10
Step 6: Implement the function to manage the elite	10
Step 7: Implement the genetic algorithm	20
Step 8: Find the problems' solution (all the solutions in the Pareto front). You need to draw the solutions in the 2d map (error and #features), and for each solution indicates what features were selected.	20