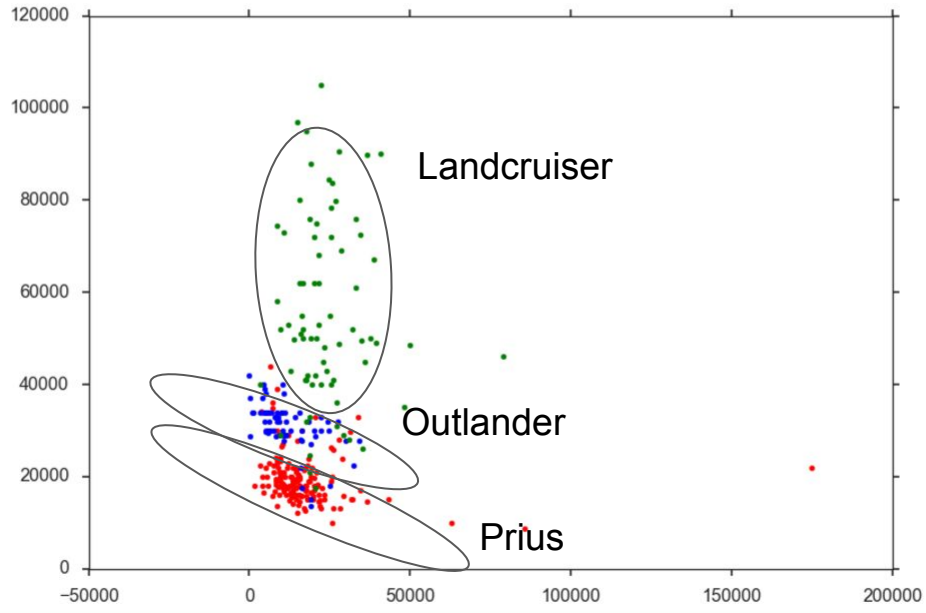
capstone project [part3]

updates

DATA. changes

```
In [16]: #prius.plot.scatter('utilization_ratio', 'price')
         matplotlib.pyplot.scatter(prius.utilization_ratio, prius.price, color = 'red')
         matplotlib.pyplot.scatter(outlander.utilization_ratio, outlander.price, color = 'blue')
         matplotlib.pyplot.scatter(Landcruiser.utilization_ratio, Landcruiser.price, color = 'green')

Out[16]: <matplotlib.collections.PathCollection at 0xc735668>
```



each model has its own rate of price decrease

in real life when you want to buy a car you know which make you want,

so let’s solve real problem.

better to stick to one particular make...

ok let’s buy most common… say Toyota Corolla

DATA changes

SCRAPING AGAIN….

…..

…..

CLEANING AGAIN

…….

clean data set with 3397 cars

	fuel_efficiency	price	year	odometer	Number of cylinders	capacity
count	3397.0	3397.000000	3397.000000	3397.000000	3397.0	3397.000000
mean	11.0	12712.374154	2009.443921	93639.964086	4.0	1.793318
std	0.0	6325.436554	5.203926	69126.368582	0.0	0.037195
min	11.0	400.000000	1984.000000	1.000000	4.0	1.300000
25%	11.0	7990.000000	2007.000000	41500.000000	4.0	1.800000
50%	11.0	12888.000000	2011.000000	76801.000000	4.0	1.800000
75%	11.0	16990.000000	2014.000000	133600.000000	4.0	1.800000
max	11.0	35449.000000	2016.000000	785500.000000	4.0	2.000000

DATA.changes

ok, Let’s buy Toyota Corolla

… oops there are 3397 Toyota Corolla cars around Australia

which one is the best deal?

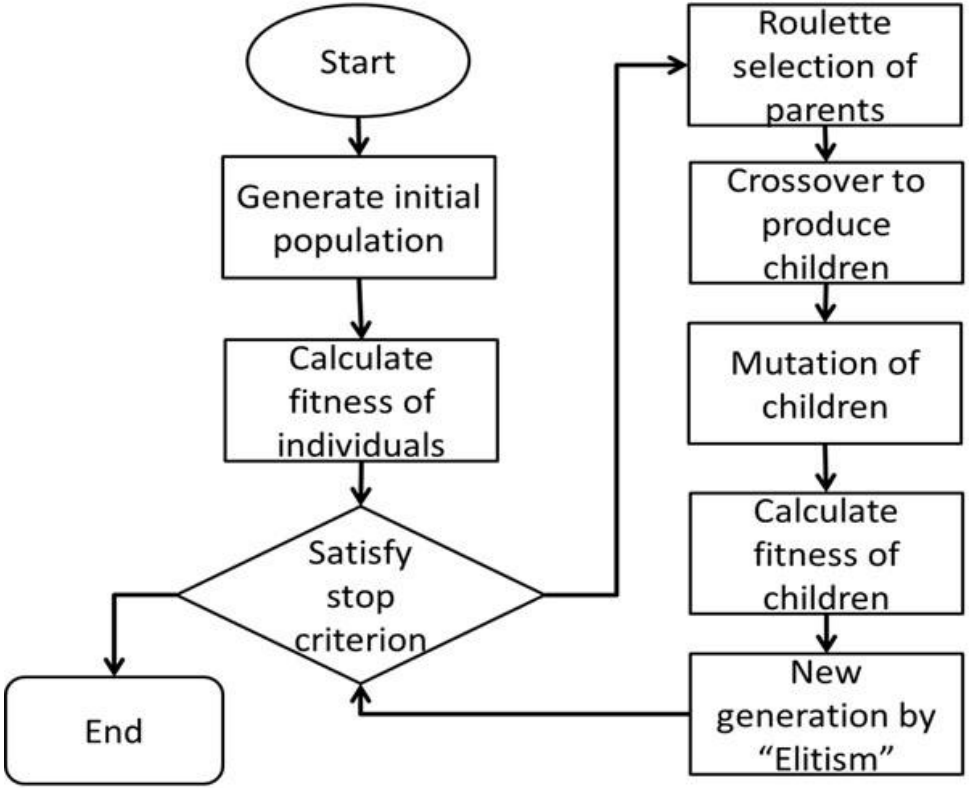
Multi-Objective Problem...

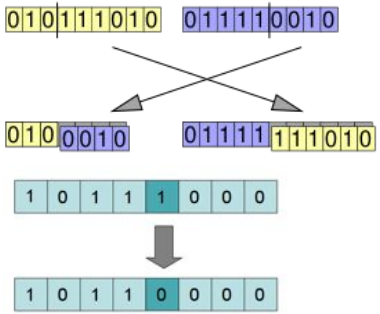
GENETIC ALGORITHMS

In computer science and operations research, a **genetic algorithm** (**GA**) is a metaheuristic inspired by the process of natural selection that belongs to the larger class of evolutionary algorithms (EA). Genetic algorithms are commonly used to **generate high-quality solutions to optimization and search problems** by relying on bio-inspired operators such as mutation, crossover and selection. [WIKI]

- A genetic algorithm is a search heuristic that mimics the process of natural evolution.
- There are five phases
  - Initial Population
  - Fitness Function
  - Selection
  - Crossover
  - Mutation
- The primary advantage of GA's comes from the crossover operation.

GENETIC ALGORITHMS





the best organism(s) from the current generation to carry over to the next, unaltered

NSGA-II. PyBRAIN library

Non-dominated Sorting Genetic Algorithm-II (**NSGA-II**)