Fundamentals of Artificial Intelligence

VICTORIA UNIVERSITY OF WELLINGTON
TE HERENGA WAKA

COMP307/AIML420

Evolutionary Computation: Tutorial

Evolutionary Computation

- Evolutionary algorithms (genetic operators, e.g., crossover)
  - Genetic algorithms (the biggest branch)
  - Evolutionary programming
  - Evolution strategies
  - Genetic Programming (Koza, 1990s, fast growing area)

- Swarm intelligence (no genetic operators)
  - Ant colony optimization
  - Particle swarm optimization (PSO)
  - Artificial immune systems

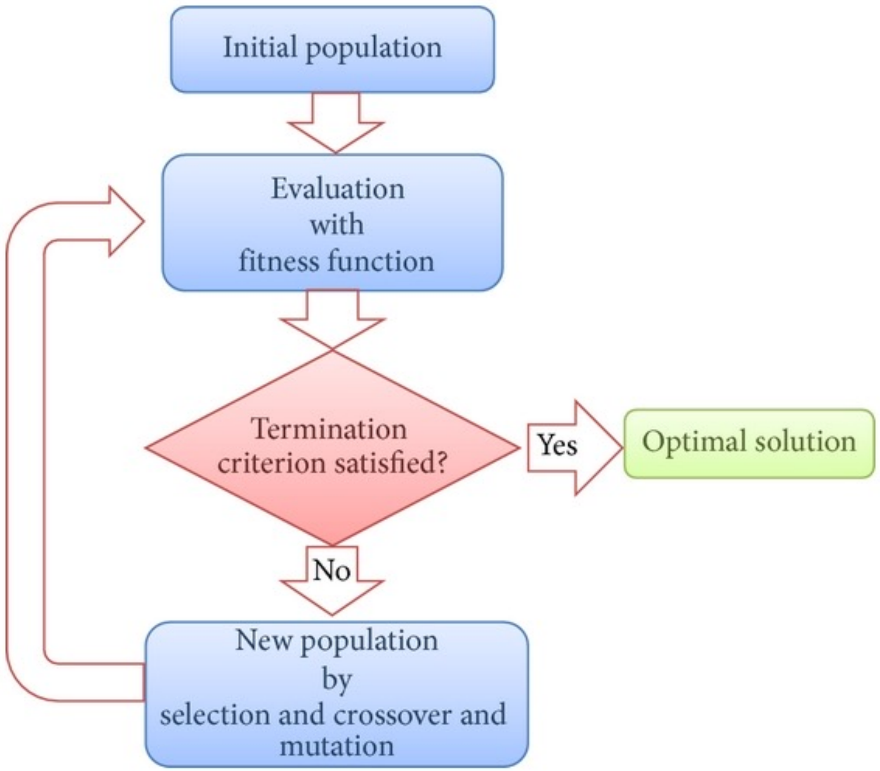
- Other techniques
  - Differential evolution
  - Estimation of distribution algorithms
  - …

Tutorial question

- Identify two main differences between EAs and swarm intelligence algorithms.

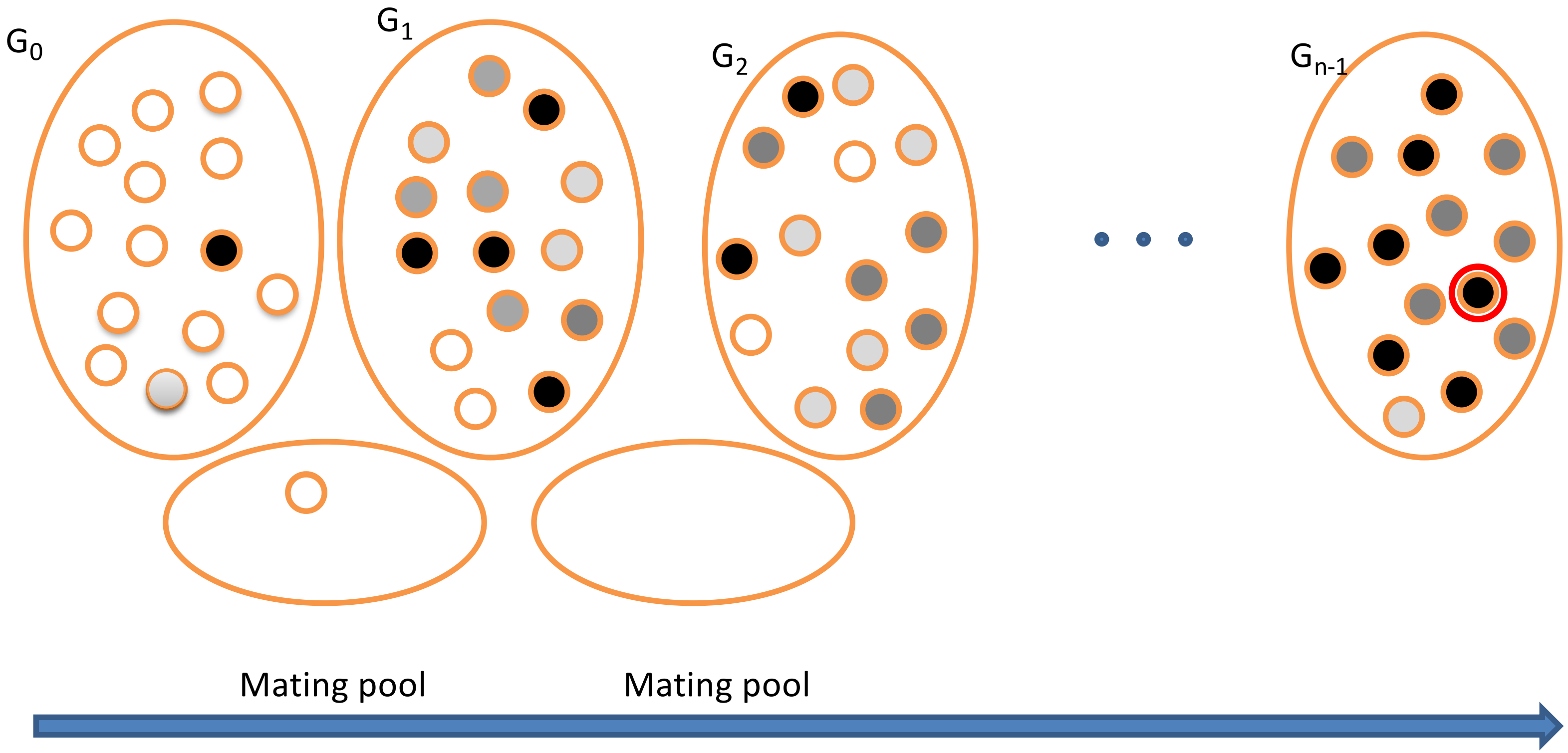
Evolutionary Computation

- Three steps of learning



Evolutionary Algorithms

- Darwinian biological evolution principle
  - Representation is problem dependent
  - Fitness function: goodness measure
  - Selection: better individuals are more likely to survive and produce offspring
  - Genetic operators: to generate new individuals (crossover, mutation)



Tutorial question

- Identify at least three stopping criteria that can be used by an evolutionary algorithm.

Tutorial question

- Given a population of four individuals as summarized in the table below, calculate the probability of selecting any individual as the parent upon using the ranked based selection method.

Individual	Fitness
Ind1	44
Ind2	27
Ind3	58
Ind4	63

A Basic Genetic Algorithm

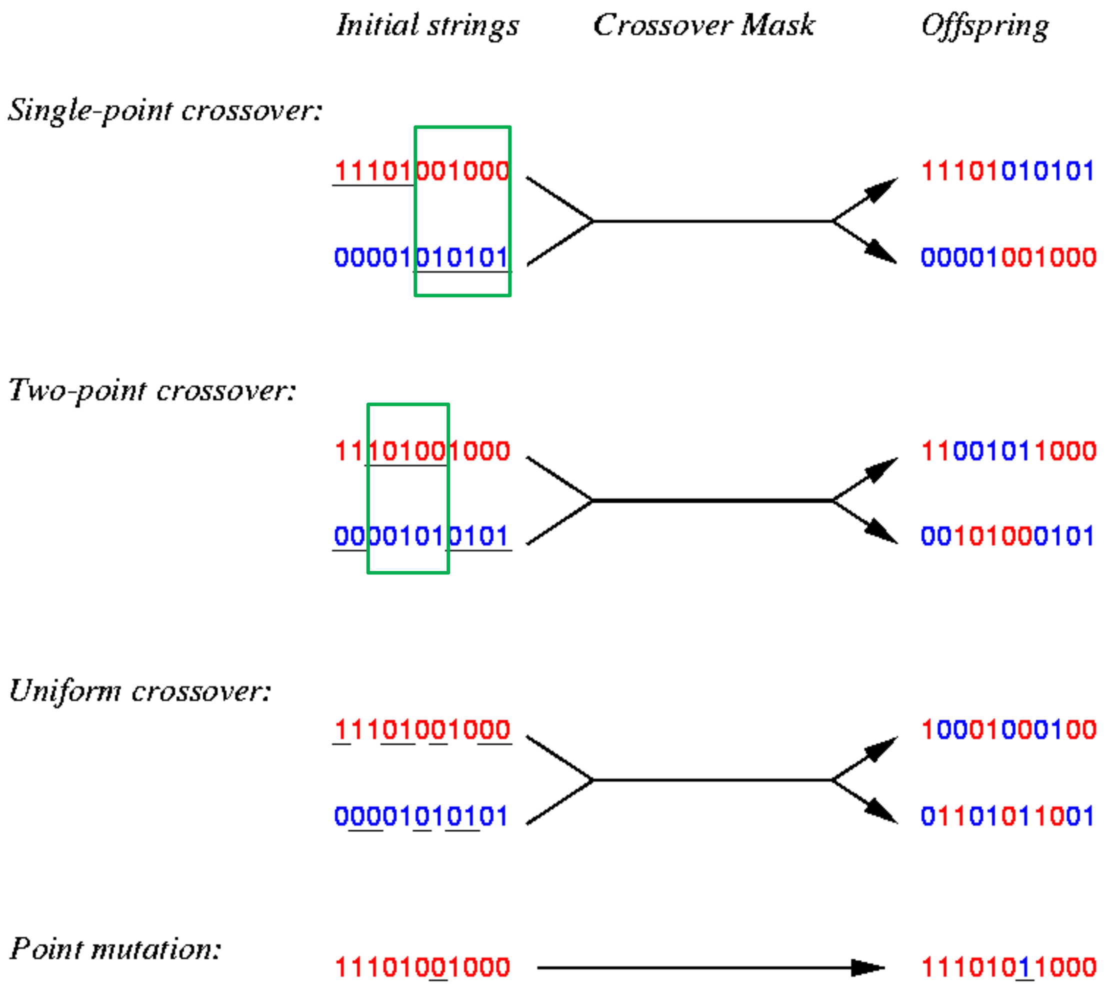
- Randomly initialise a population of chromosomes

- **Repeat until** stopping criteria are met:
  - Construct an empty new population
  - **Repeat until** the new population is full:
    - Select **two** parents from the population by roulette wheel selection
    - Apply crossover to the two parents to generate two children
    - Each child has a probability (mutation rate) to undergo mutation
    - Put the two children into the new population
  - **End Repeat**
  - Move to the new population (new generation)
- **End Repeat**

- Output the best individual from the final population

Genetic Algorithm

- Representation: individuals are binary strings
- An individual is also called a chromosome



Tutorial question

- Suppose every individual in GA is representation as a permutation of integers 1 to 5, propose a crossover and a mutation operator that can work well with this solution representation?

A Simple GA Example

- **OneMax Problem**

to find the binary string of a given length that maximizes the sum of its digits.

  - Target to (11111…1)
  - More zeros means worse: far away from the target
  - Simple “benchmark” problem!

- Representation: bit string
- Fitness function: $1 + \sum x_i$ (the larger, the better)
  Or just $\sum x_i$ (more greedy), Or $100 + \sum x_i$ (any constant number)
- Crossover: single-point crossover
- Mutation: single-point mutation

A Simple GA Example

- 10 bits (Optimal fitness = 11)
- population size = 20
- mutation rate = 0.1 (10%), crossover rate = 0.8 (80%), reproduction rate = 0.1 (10%)
- Run for 10 generations

```
At generation 0 average fitness is 6.0, best fitness is 9
At generation 1 average fitness is 6.65, best fitness is 10
At generation 2 average fitness is 6.8, best fitness is 11
At generation 3 average fitness is 6.9, best fitness is 9
At generation 4 average fitness is 6.45, best fitness is 9
At generation 5 average fitness is 6.95, best fitness is 9
At generation 6 average fitness is 7.3, best fitness is 11
At generation 7 average fitness is 6.65, best fitness is 10
At generation 8 average fitness is 6.25, best fitness is 8
At generation 9 average fitness is 6.6, best fitness is 8
```

Keep elites (i.e., best ones) to the next generation!!!

Code in Python

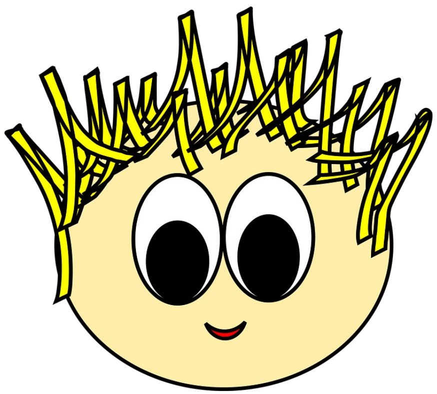
- EC is a stochastic algorithm
- Need a seed to generate the random numbers for repeating results
- 30 runs (or more) with statistical test to measure performance

- https://github.com/DEAP/deap/blob/master/examples/ga/onemax.py

GP for Symbolic Regression

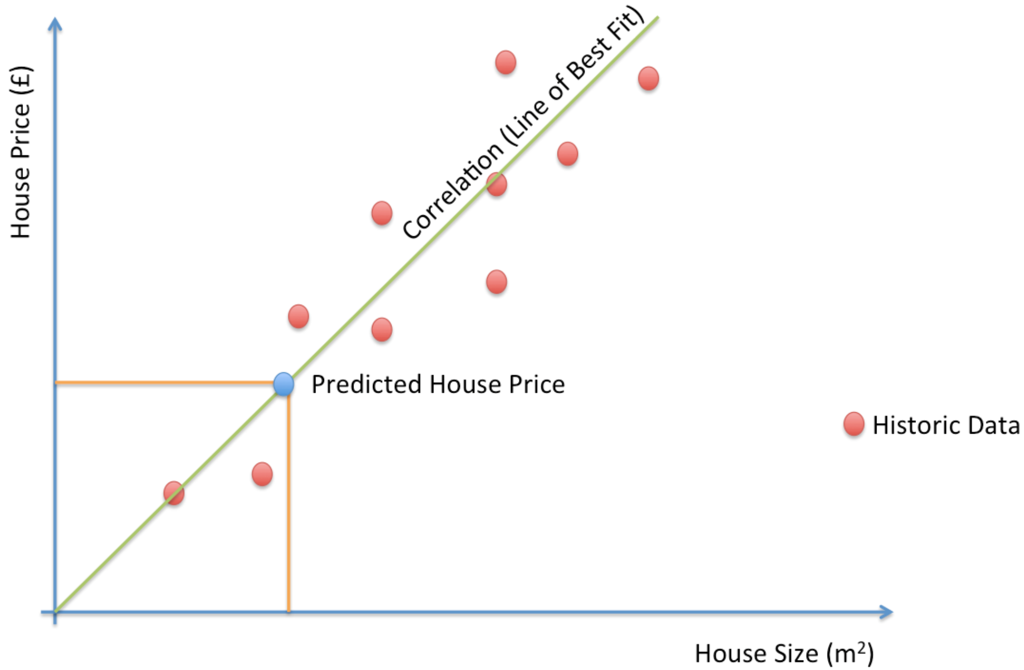
- Statistical (parameter) regression
- Symbolic regression
- GP for symbolic regression

Learn the pattern of data



House Price Prediction

- How much should your tender be?



Price	Floor space	Rooms	Lot size
250000	71	4	92
209500	98	5	123
349500	128	6	114
250000	86	4	98
419000	173	6	99
225000	83	4	67
549500	165	6	110
240000	71	4	78
340000	116	6	115

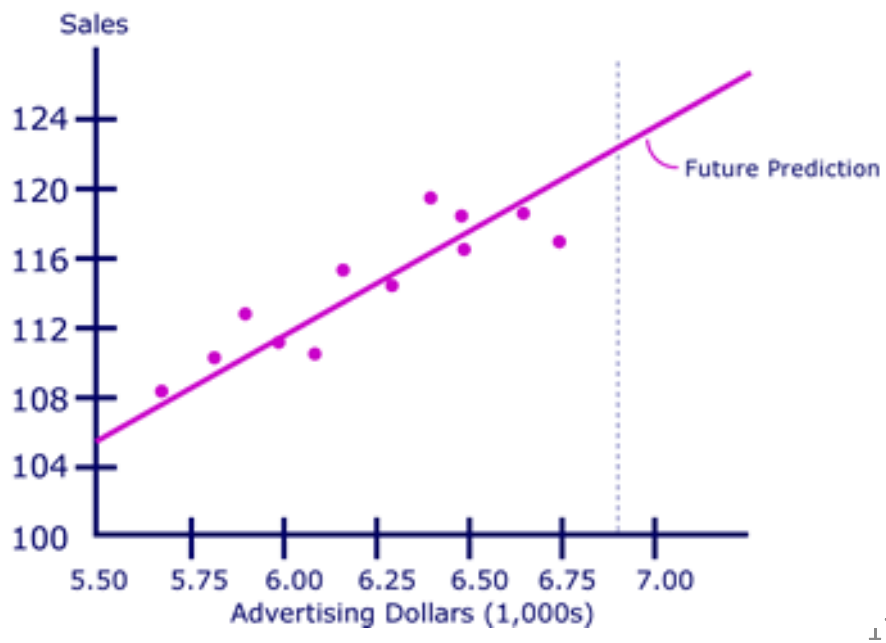
(Statistical) Regression Analysis

- In statistics, regression analysis examines the relation of a dependent variable (*response variable, y*) to specified independent variables (*explanatory variables, x*)

  $y = x^2$

  - The mathematical/regression model of their relationship is the regression equation (e.g. $f(x, y) = 0$), e.g., $y - x^2 = 0$
  - estimates of one or more hypothesized regression parameters ("constants")

- Examples
  - Financial prediction
  - Sales prediction
  - Ad cost vs sales

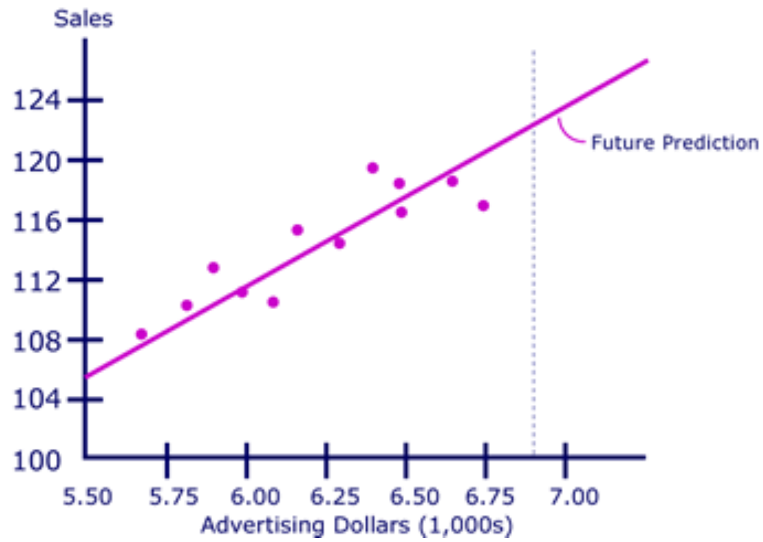


(Statistical) Regression

- Two main tasks:
  - **Regression equation**: relationship between a dependent variable (response variable) and independent variables (explanatory variables)
  - **Parameters/Coefficients**

- Example: linear regression
  - Regression equation is a linear model: $y = \alpha \cdot x + \beta + \epsilon$
  - Coefficients: α and β
  - α is the slope, β is the intercept
  - ϵ is the *error* term (assume normally distributed)
  - Estimate α and β to minimise ϵ

$$\epsilon = y - \alpha \cdot x - \beta$$

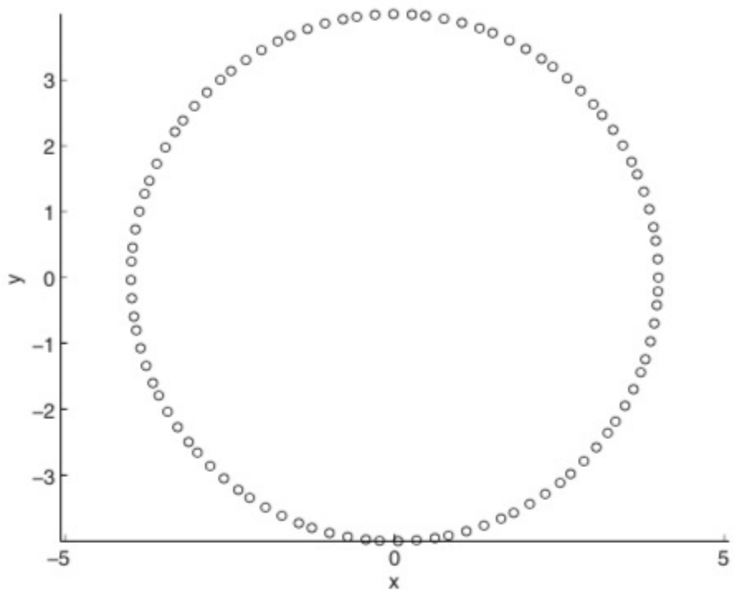


Assume model structure, estimate model parameters

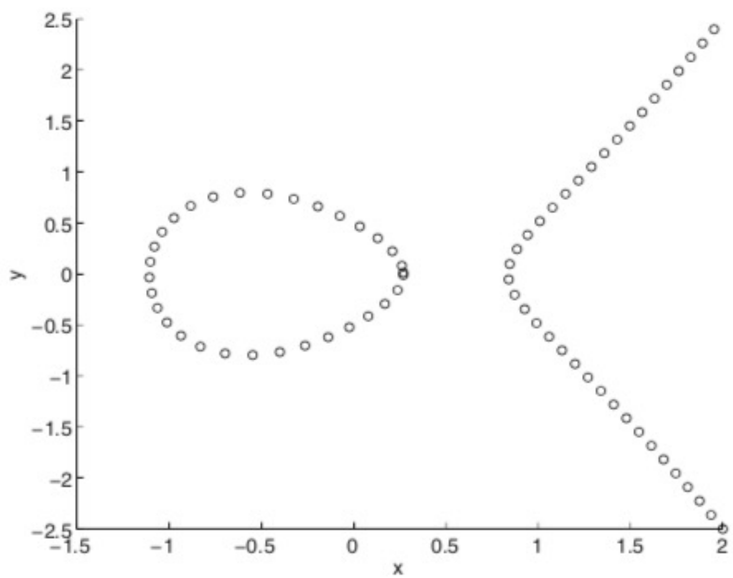
Symbolic Regression

- However, linear model is too simple for many real-world data

- Hard to find the proper regression equation



(a) **Circle:** $x^2 + y^2 - 4^2$



(b) **Elliptic Curve:** $x^3 + x - y^2 - 1.5$

Symbolic Regression Applications

- Symbolic regression has many real-world applications:
  - Economics prediction: stock market prediction, GDP prediction
  - Industrial prediction: prediction of container handling capacity at a sea port; short/medium/long-term prediction of power load at a region
  - Experimental formula modelling in engineering: formulating the amount of gas emitted from a coal surface
  - Time series projection: Gross National Happiness (GNH)
  - **Others?**

Define Symbolic Regression

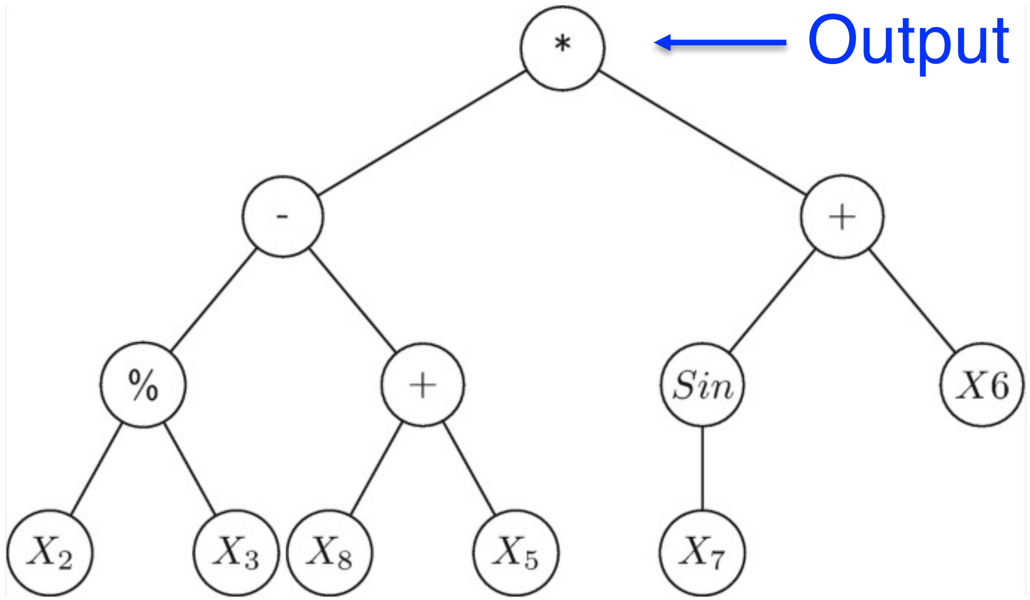
- Symbolic regression: finds a symbolic description of a model, not just a set of coefficients/parameters in a pre-specified model.

- Finds both:
  - the model structure, and
  - the corresponding coefficients/parameters

GP for Symbolic Regression (Example)

- Objective: Find a program/model that produces the correct value of the dependent variable y when given the value of an independent variable x
- Terminal Set: *x*, random constant
- Function Set: $\{+, -, \times, \div$ (protected)$\}$
- Fitness Cases: 50 cases of *x* and the corresponding *y* values (e.g. *50 instances/patterns/cases*)
- Fitness Measure: Sum/Average of the *absolute* errors for the 50 cases
- Parameters: Population = 100, Generations = 50, MaxDepth = 17 reproduction rate: 5%, crossover rate: 90%, mutation rate: 5%
- Success: The fitness value is smaller than a pre-defined value, e.g., 0.01
- Termination criteria: satisfactory solutions found, or at generation 50

Regression Model (example)



- Assume there are 8 features and this regression model has the best fitness
- All features important?

- Not all the features are important
- GP can automatically do feature selection (to some extent)

Code in Python

- Check the linked code for an example of using GP for symbolic regression

- https://github.com/DEAP/deap/blob/master/examples/gp/symbreg.py

Tutorial question

- Genetic Programming (GP) is considered a good method for binary classification. Consider using the standard tree-based GP to solve a binary classification problem with the following training set (X1 – X5 are the features),

X1	X2	X3	X4	X5	Class
1.2	3.1	9.1	11.3	12.0	O
4.1	7.3	14.2	22.1	2.1	X
0.6	5.3	11.1	13.2	9.8	O
1.5	1.3	7.8	14.1	15.6	O
1.2	8.8	5.5	19.1	7.5	X

- Choose a good terminal set.
- Choose a good function set.
- Describe how to use evolved GP tree to make classification decisions.