Machine Learning Using Karoo GP

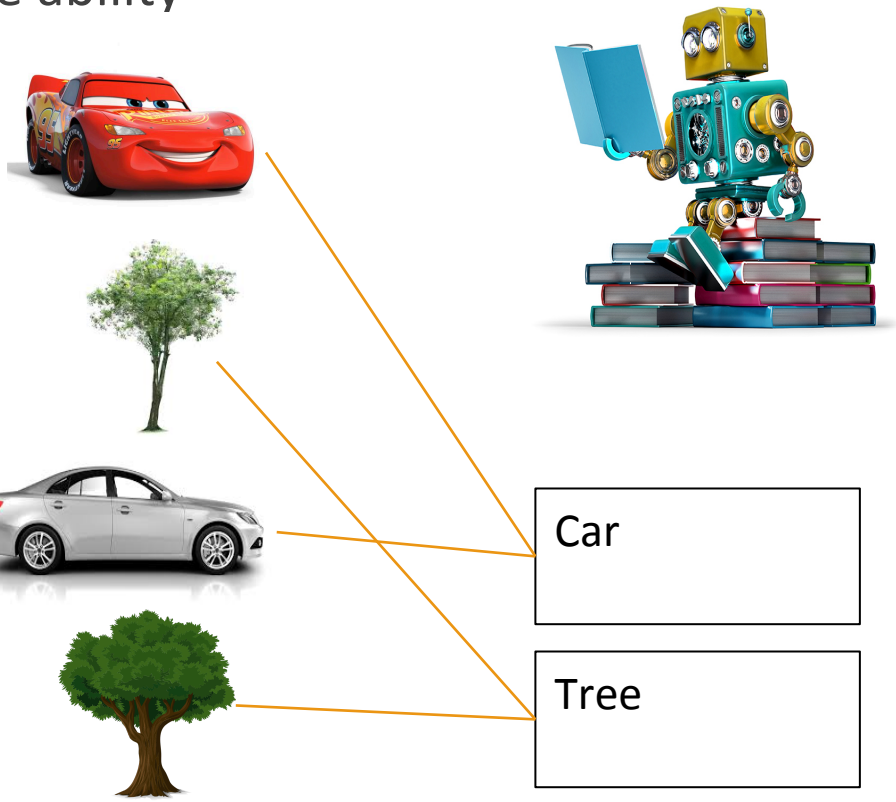
ASPIRE WORKSHOP

SUMMER 2017

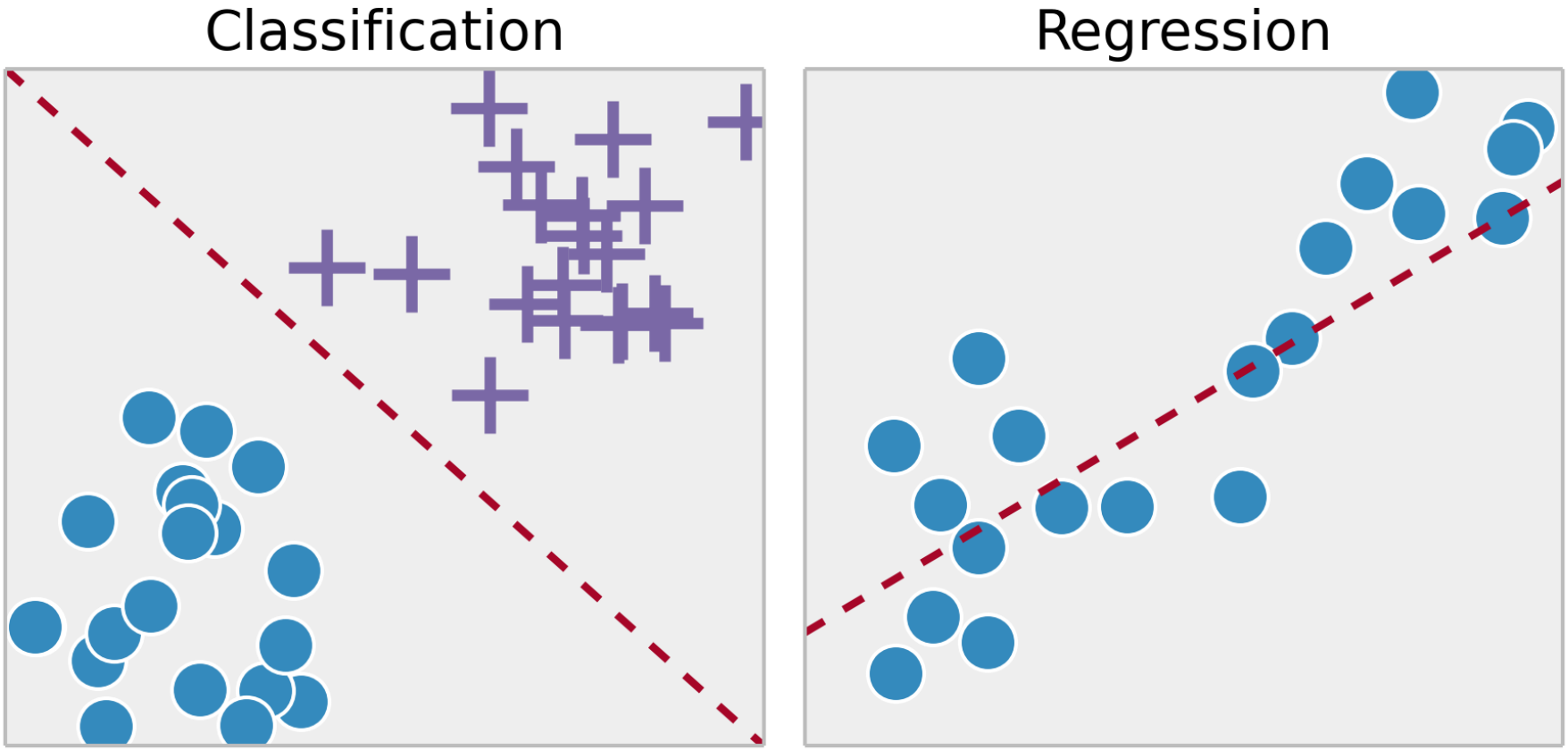
THE OHIO STATE UNIVERSITY

What is machine learning?

- Machine learning (ML) gives computers the ability to learn without explicit programming
- The goal is to find a structure that maps a given input to a desired output
- Examples of problems ML could solve:
  - Netflix movie recommendations
  - Categorizing images
  - Online advertising
  - Fitting data to a function



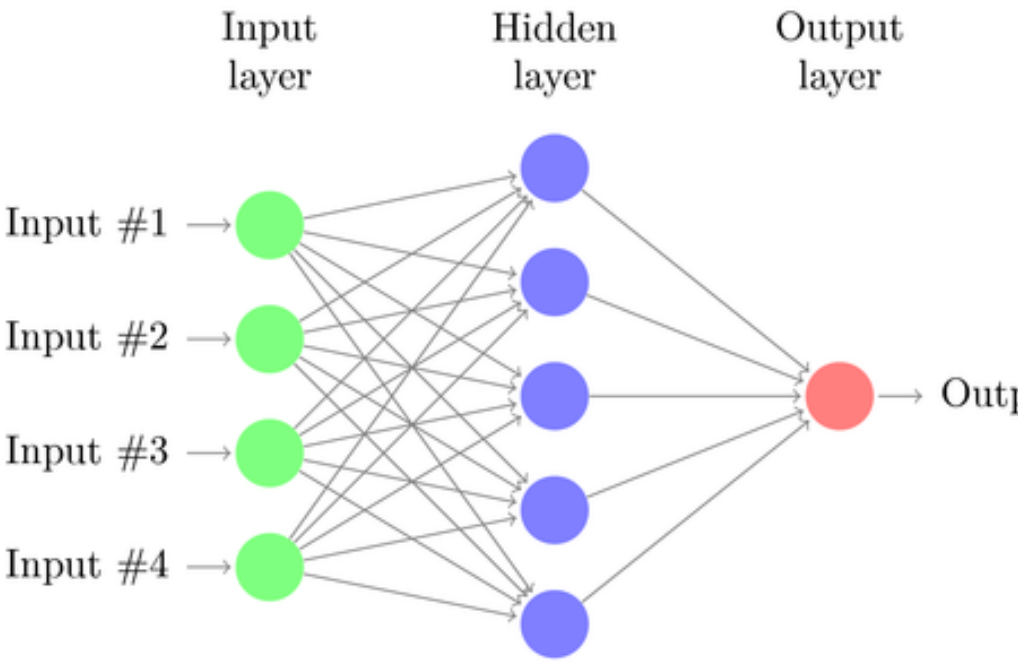
Regression vs. Classification



Some approaches to machine learning

Artificial Neural Networks:

- They work like neurons in our brains!
- The hidden layer uses a series of operations to take a set of inputs and turn it into the desired output, similar to how a brain processes data



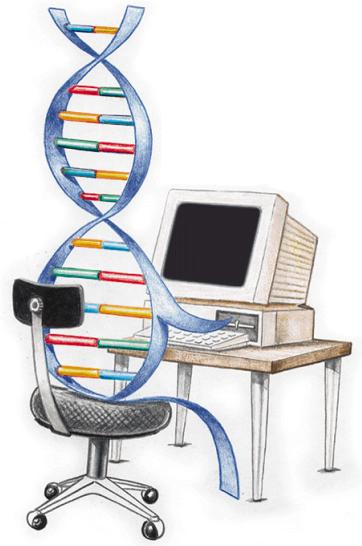
Genetic Programming (GP)

What would happen if we applied evolution to a Regression Problem?

Genetic programming approach:

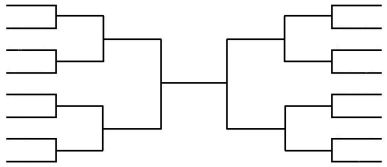
- Create a group of random solutions

  Ex: $1\uparrow st$ solution $=2x\uparrow 3+e\uparrow y$, $2\uparrow nd$ solution$=\cos(x)/x$, $etc.$
- Test and evaluate each solution
- Keep the best solutions and “kill" the rest
- Mutate the best solutions to produce even better ones

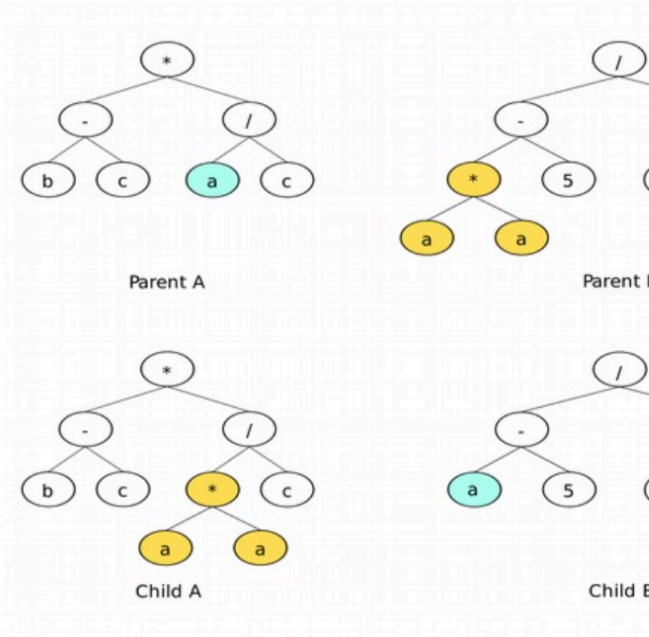


Genetic Programming (Cont.)

- Each solution is tested based on a fitness function
  - For Regression: Uses a least squares method
  - For Classification: Uses a relative absolute error



- The best solutions are chosen through a tournament selection
- The mutation process works like the following:
  - Reproduction: The function is copied to the next generation
  - Point Mutation: One thing is change ex: 2*t/(x+y) to 2*t/(x+t)
  - Branch Mutation: A full branch is changed ex: 2*t/(x+y) to 2*t/(y^2*t)
  - Crossover: 2 parents combine to produce offspring (see picture)



Karoo GP

- Kai Staats created a GP software called Karoo GP
- Karoo was Kai's master's thesis at the University of Cape Town
- Karoo is used to mitigate backgrounds for the LIGO gravitational wave observatory
- In addition to Genetic Programming Kai also enjoys rock climbing, making documentaries, creating radio telescopes and attending Space University

Karoo GP Cont.

Karoo GP is used to perform various tasks:

- Regression: Find a relationship between two or more variables
  Ex: s = $2x^2 + 5y + z^4$
- Classification: Assign variables to different categories
  Ex: Distinguish between pictures of roses, sunflowers and a peonies
- Matching: Find a solution that matches the output with the inputs
  Ex: a+b+c = s.