Evolutionary Computation

Mimicking nature’s optimizer

Guilherme Seidyo Imai Aldeia

November 18, 2024

Who designed them?

Who designed them?

Nature did!

Evolution: nature’s optimizer

It is the combination of **Genetic variation** and **Survival of the fittest**:

- If you are better suited for the environment, you are more likely to survive;
- The offspring will inherit the parent’s characteristics;
- There is random variation when duplicating things.

Evolution: nature’s optimizer

It is the combination of **Genetic variation** and **Survival of the fittest**:

- If you are better suited for the environment, you are more likely to survive;
- The offspring will inherit the parent’s characteristics;
- There is random variation when duplicating things.

There is no optimum design. There is no right or wrong.

They are *just* better suited to the environment they exist in.

Let’s steal it!

We need:

- **Individuals**: a computational representation that we can manipulate;
- **Fitness function**: a way to evaluate how good individuals are;
- **Variation**: a way to change and recombine parts of individuals.

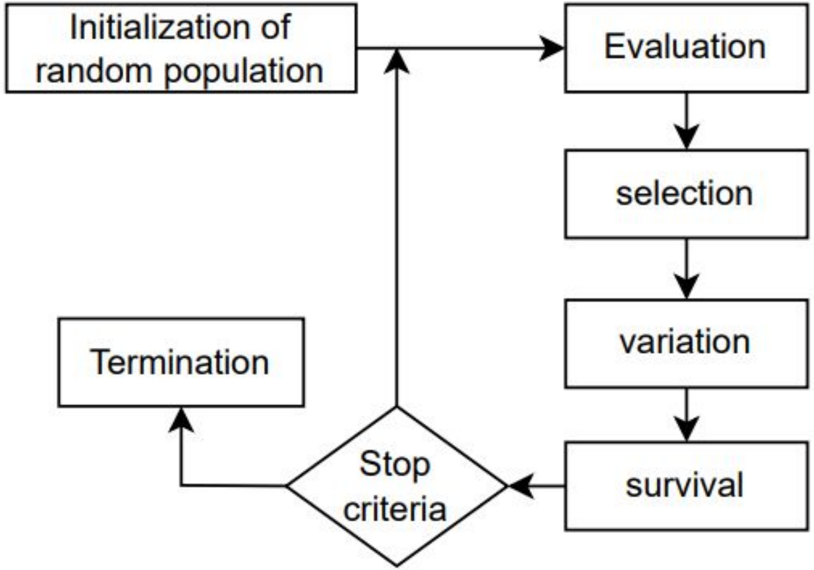
Let’s steal it!

We need:

- **Individuals**: a computational representation that we can manipulate;
- **Fitness function**: a way to evaluate how good individuals are;
- **Variation**: a way to change and recombine parts of individuals.

Now we can find good solutions without having to write (explicitly) any solution at all!

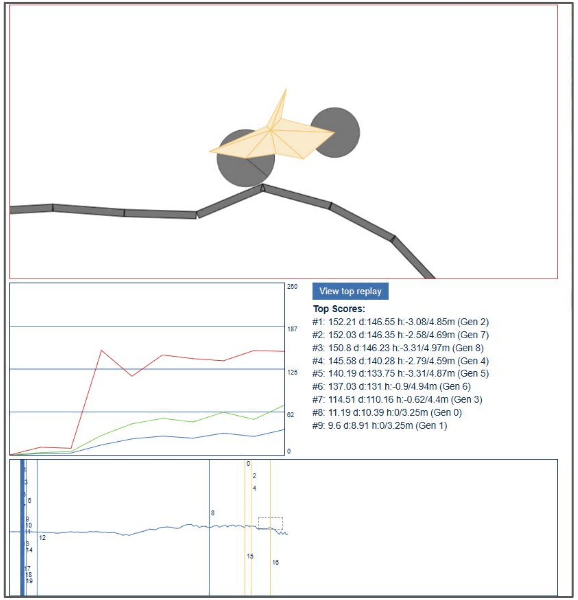
Evolutionary Algorithms



EAs are a weak metaphor for evolution, a population-based approach to *find good solutions*.

Some cool examples

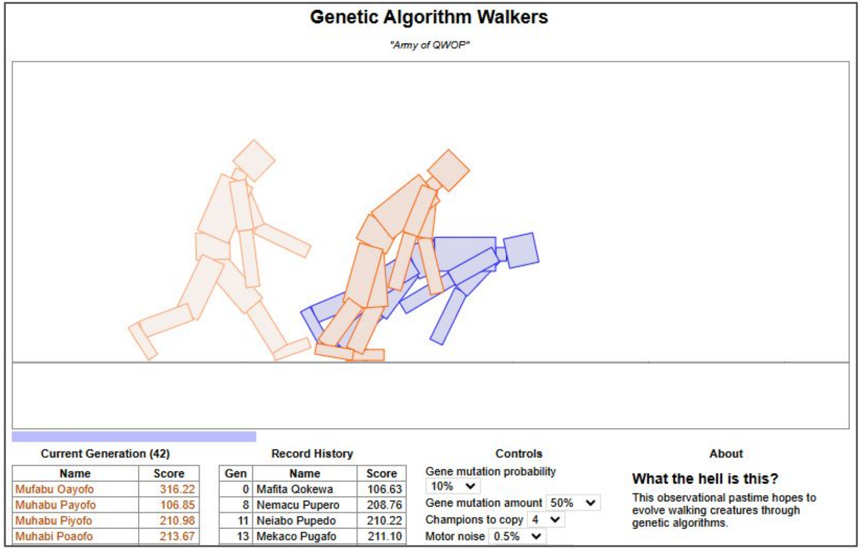
https://rednuht.org/genetic_cars_2/



- **Individuals**:
  - Geometric shapes with two wheels attached to it;
- **Fitness function**:
  - Distance travelled;
- **Variation**:
  - Mutation (*random perturbations*) on the geometry

Some cool examples

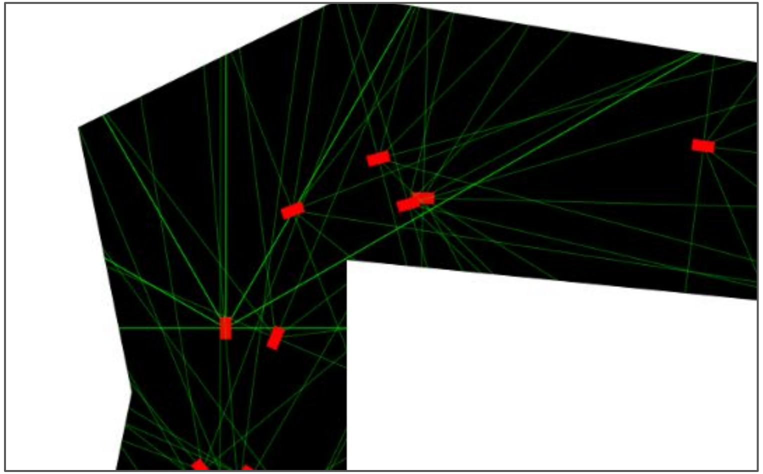
https://rednuht.org/genetic_walkers/



- **Individuals**:
  - Sequence of movements
- **Fitness function**:
  - Distance between head and ground, number of steps
- **Variation**:
  - Mutation (*random perturbations*) over movements

Some cool examples

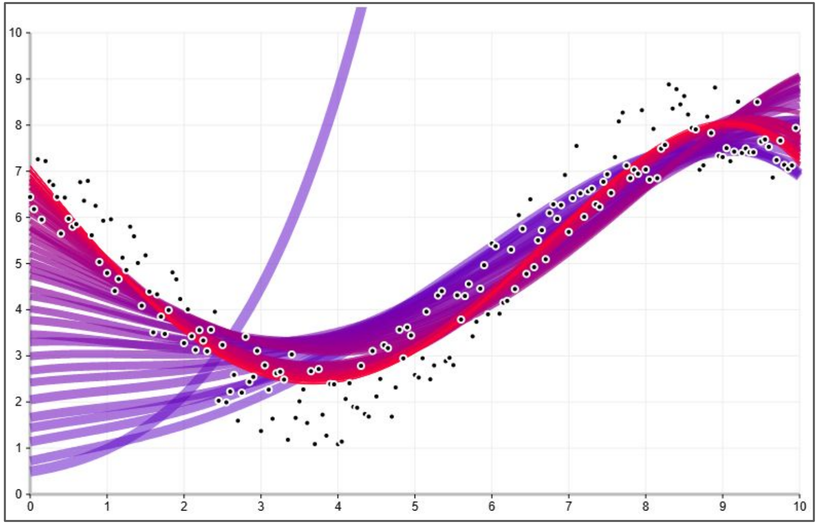
https://thunderinfy.github.io/Self-driving-2d/



- **Individuals**:
  - Thresholds of sensor readings and actions;
- **Fitness function**:
  - Distance travelled;
- **Variation**:
  - Recombination of two parents (*crossover*) and mutation;

Some cool examples

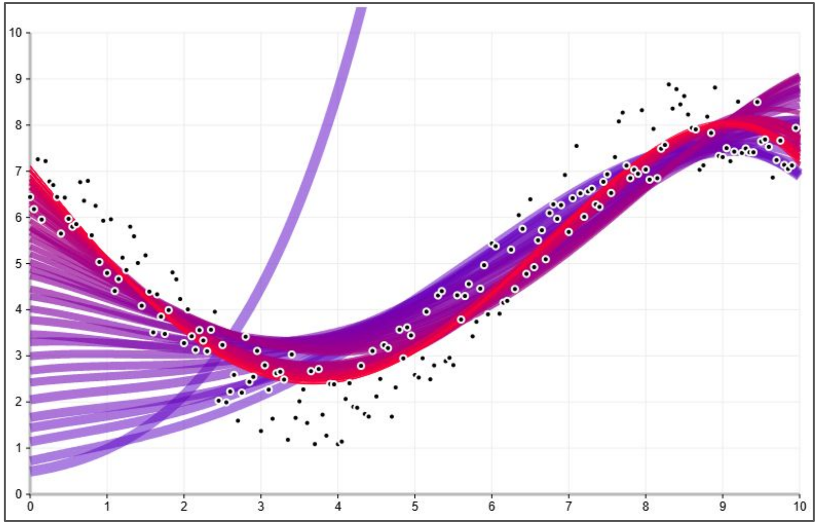
https://subprotocol.com/system/genetic-regression-curve.html



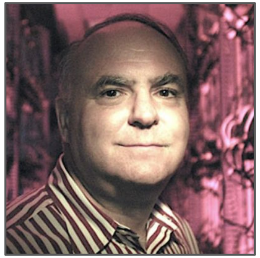
- **Individuals**:
  - Mathematical functions;
- **Fitness function**:
  - Mean error between function and points;
- **Variation**:
  - Recombination of expressions (*crossover)* and random changes of **symbols.**

Some cool examples

https://subprotocol.com/system/genetic-regression-curve.html



- **Individuals**:
  - Mathematical functions;
- **Fitness function**:
  - Mean error between function and points;
- **Variation**:
  - Recombination of expressions (*crossover)* and random changes of **symbols**.



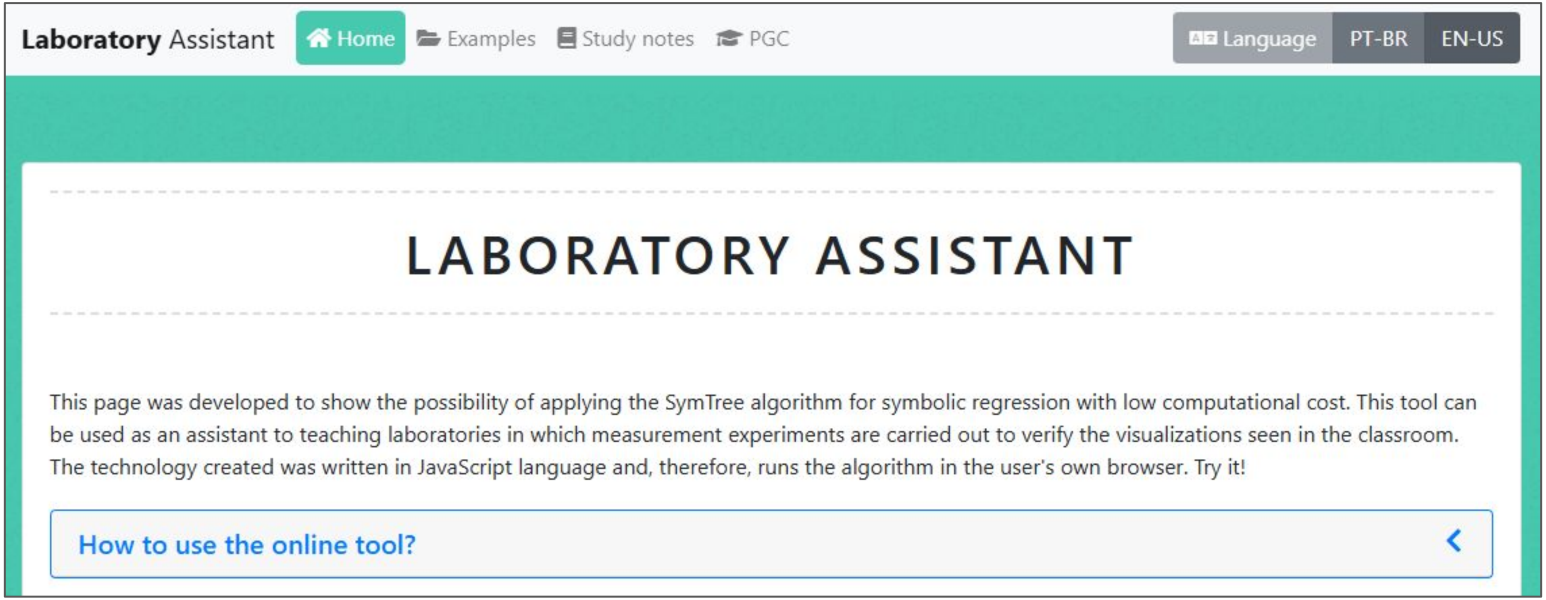
John R. Koza

Symbolic Regression

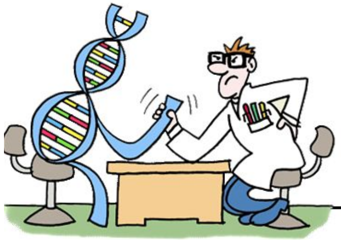
Hands-on! Let's play with **Symbolic Regression!**

https://galdeia.github.io/index-en

https://gist.github.com/gAldeia/8a1c754a01e5607b830efdd51e21098b#file-ideal_gas-csv



What else can we do?



Annual "Humies" Awards
For Human-Competitive Results

Produced By Genetic And Evolutionary Computation

Artificial Intelligence Review (2024) 57:2
https://doi.org/10.1007/s10462-023-10622-0

Check for updates

Interpretable scientific discovery with symbolic regression: a review

Nour Makke[1] · Sanjay Chawla[1]

Accepted: 1 October 2023 / Published online: 2 January 2024

Abstract
Symbolic regression is emerging as a promising machine learning method for learning succinct underlying interpretable mathematical expressions directly from data. Whereas it has been traditionally tackled with genetic programming, it has recently gained a growing interest in deep learning as a data-driven model discovery tool, achieving significant advances in various application domains ranging from fundamental to applied sciences. In this survey, we present a structured and comprehensive overview of symbolic regression methods, review the adoption of these methods for model discovery in various areas, and assess their effectiveness. We have also grouped state-of-the-art symbolic regression applications in a categorized manner in a living review.

Archives of Computational Methods in Engineering (2023) 30:3845–3865
https://doi.org/10.1007/s11831-023-09922-z

REVIEW ARTICLE

Check for updates

Artificial Intelligence in Physical Sciences: Symbolic Regression Trends and Perspectives

Dimitrios Angelis[1] · Filippos Sofos[1] · Theodoros E. Karakasidis[1]

Received: 4 December 2022 / Accepted: 27 March 2023 / Published online: 19 April 2023

Abstract
Symbolic regression (SR) is a machine learning-based regression method based on genetic programming principles that integrates techniques and processes from heterogeneous scientific fields and is capable of providing analytical equations purely from data. This remarkable characteristic diminishes the need to incorporate prior knowledge about the investigated system. SR can spot profound and elucidate ambiguous relations that can be generalizable, applicable, explainable and span over most scientific, technological, economical, and social principles. In this review, current state of the art is documented, technical and physical characteristics of SR are presented, the available programming techniques are investigated, fields of application are explored, and future perspectives are discussed.

What else can we do?

npj | digital medicine

www.nature.com/npjdigitalmed

ARTICLE OPEN

Check for updates

A flexible symbolic regression method for constructing interpretable clinical prediction models

William G. La Cava[1,5], Paul C. Lee[2,5], Imran Ajmal[2], Xiruo Ding[2], Priyanka Solanki[2], Jordana B. Cohen[3,4], Jason H. Moore[4] and Daniel S. Herman[2]✉

Machine learning (ML) models trained for triggering clinical decision support (CDS) are typically either accurate or interpretable but not both. Scaling CDS to the panoply of clinical use cases while mitigating risks to patients will require many ML models be intuitively interpretable for clinicians. To this end, we adapted a symbolic regression method, coined the feature engineering automation tool (FEAT), to train concise and accurate models from high-dimensional electronic health record (EHR) data. We first present an in-depth application of FEAT to classify hypertension, hypertension with unexplained hypokalemia, and apparent treatment-resistant hypertension (aTRH) using EHR data for 1200 subjects receiving longitudinal care in a large healthcare system. FEAT models trained to predict phenotypes adjudicated by chart review had equivalent or higher discriminative performance ($p < 0.001$) and were at least three times smaller ($p < 1 \times 10^{-6}$) than other potentially interpretable models. For aTRH, FEAT generated a six-feature, highly discriminative (positive predictive value = 0.70, sensitivity = 0.62), and clinically intuitive model. To assess the generalizability of the approach, we tested FEAT on 25 benchmark clinical phenotyping tasks using the MIMIC-III critical care database. Under comparable dimensionality constraints, FEAT's models exhibited higher area under the receiver-operating curve scores than penalized linear models across tasks ($p < 6 \times 10^{-6}$). In summary, FEAT can train EHR prediction models that are both intuitively interpretable and accurate, which should facilitate safe and effective scaling of ML-triggered CDS to the panoply of potential clinical use cases and healthcare practices.

npj Digital Medicine (2023)6:107 ; https://doi.org/10.1038/s41746-023-00833-8

Thank you!

Guilherme Aldeia

guilherme.aldeia@ufabc.edu.br

https://github.com/gAldeia

Evolutionary Computation

Mimicking nature’s optimizer

Spare questions (in case someone needs it)

1. What do you need to solve a problem with EAs?
2. What if the final solution is not what you expected?
3. Are we living in a simulation?
4. Can you guide the evolution?
5. Do we have more sophisticated EAs?
6. Do biologists like this idea?
7. What are some examples of research topics in the field of EAs?