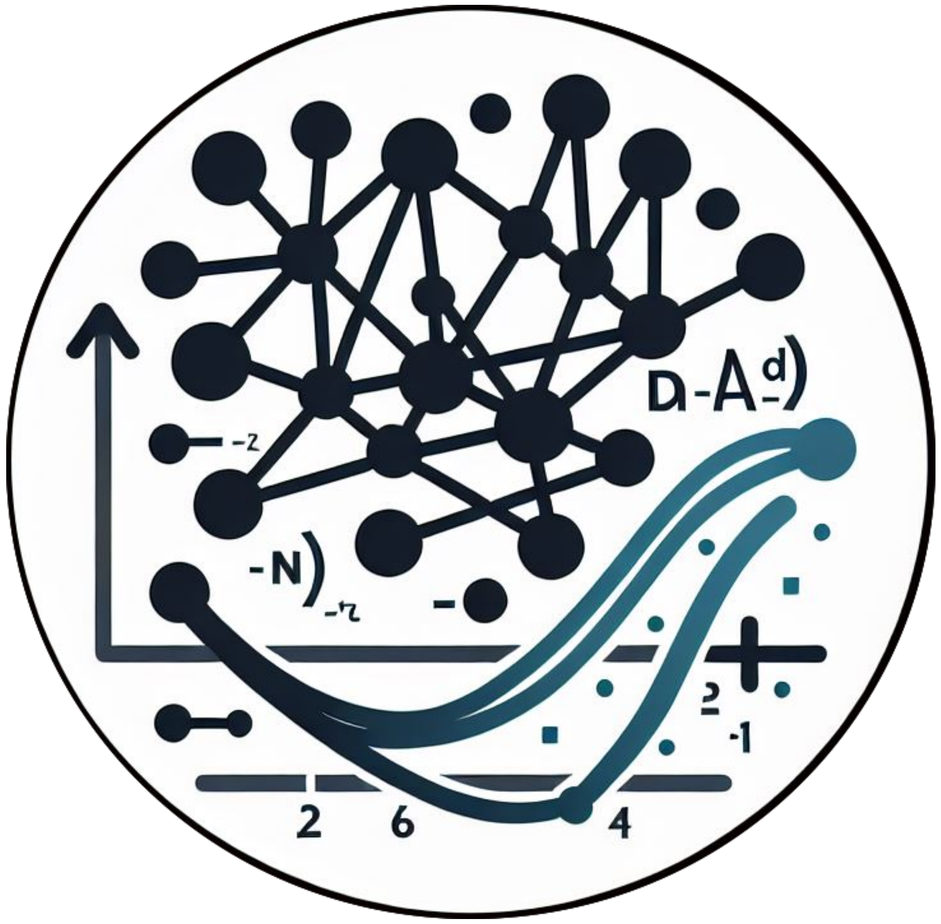
Symbolic Regression



Deep Learning for Engineers

Andrew Ning
aning@byu.edu

Project

Proposal (next week's HW): meet with me or one of the TAs to checkoff your idea

Timesheet (~70% project score). 24 hours.

Presentation (~20%)

Watch/Review Others (~10%)

Mini-Review

Multilayer perceptron (MLP): starting point for tabular data or a building block in larger architectures.

Physics-informed neural network (PINN): Embed PDE/ODE residuals in the loss function. Enhance accuracy, particularly when data is limited.

Neural ODE: Parameterize $dy/dt = f(y,t)$ with a neural network. Use regular ODE solvers. Can learn unknown dynamics, or speedup known dynamics.

Autoencoder (AE): change dimensionality (lift to higher dimensions, or compress to lower dimensions). Primarily used in conjunction with other methods.

Neural Koopman: like neural ODE, but learn *linear* dynamics. Can take advantage of convergence analysis, linear control theory, etc.

Convolutional neural network (CNN): useful when data lives on a regular grid.

Graph neural network (GNN): useful when data lives on an irregular mesh, or has varying connectivity.

Kolmogorov Arnold network (KAN): particularly useful when you need interpretability - get extract symbolic expressions.

Neural operators: learn a mapping between function spaces rather than between data to data (FNO for regular grids, DeepONets for more fleible input geometry)

Recurrent networks (RNN/LSTM/GRU): useful for sequential data (that doesn't have a governing law like an ODE)

Transformer/attention: more complex version of RNNs, especially important for long-range dependencies.

Symbolic Regression

Level 1: $y = ax^b + c\exp(d \cdot x)$

minimize $\|\hat{y} - y\|^2$

Level 2: $y = ax + bx^2 + cx^3 + \ldots + dxy + \ldots + e\sin(x) + \ldots$

minimize $\|\hat{y} - y\|^2 + \gamma\|\theta\|_1$

Level 3: structure itself is a variable

PySR and SymbolicRegression.jl

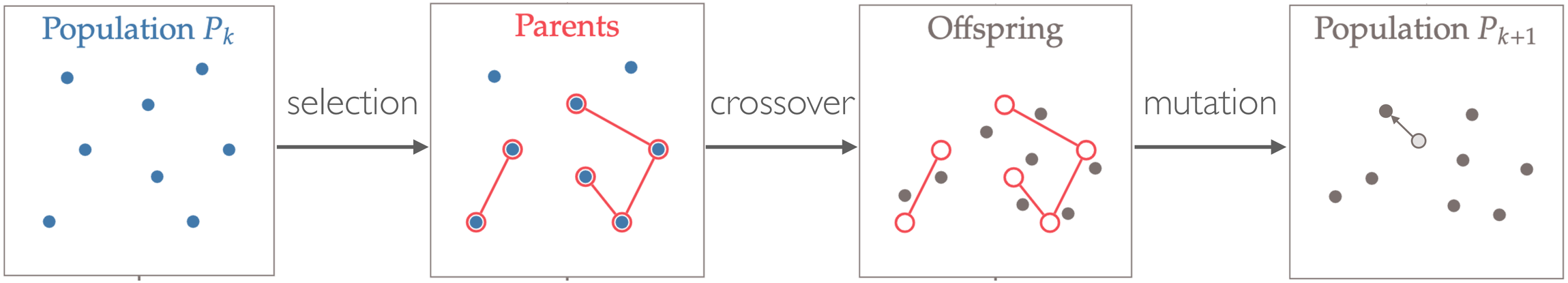
Interpretable Machine Learning for Science with PySR and SymbolicRegression.jl

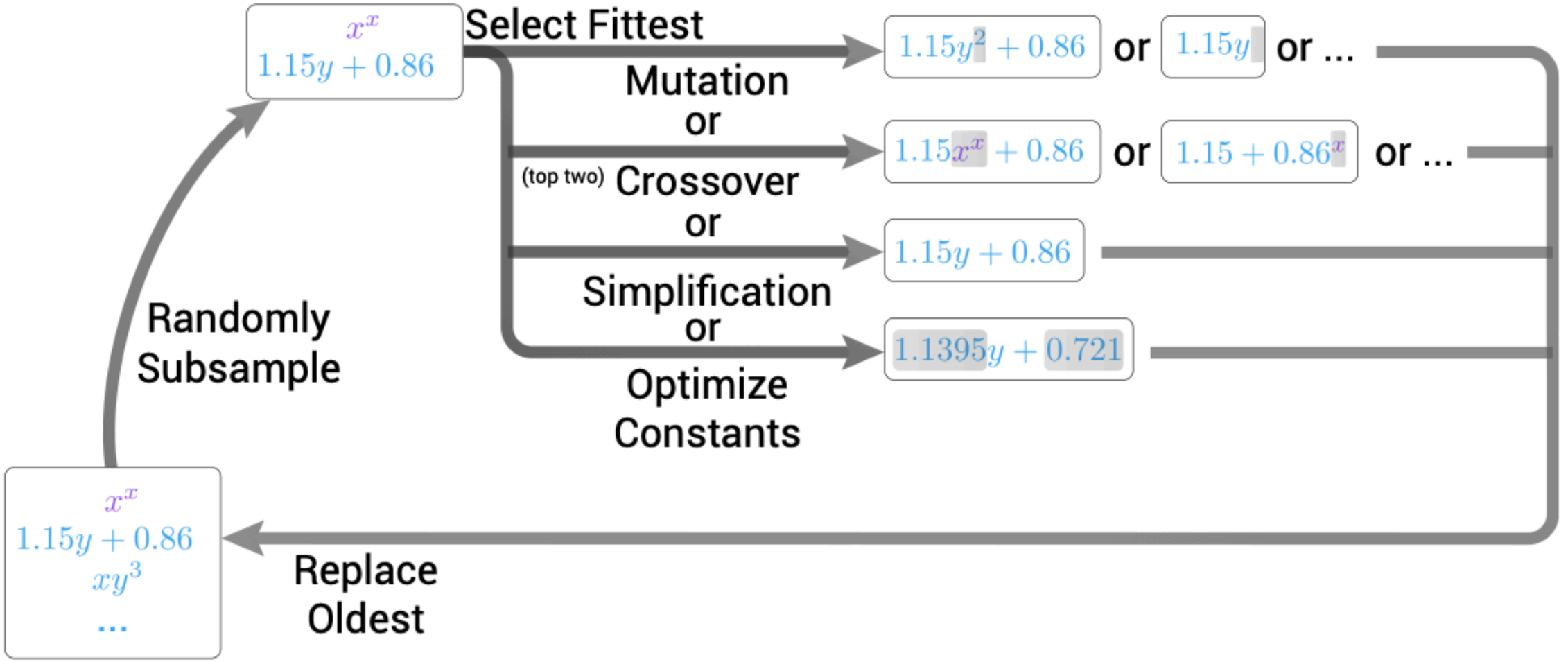
Miles Cranmer[1,2]

[1] *Princeton University, Princeton, NJ, USA*
[2] *Flatiron Institute, New York, NY, USA*

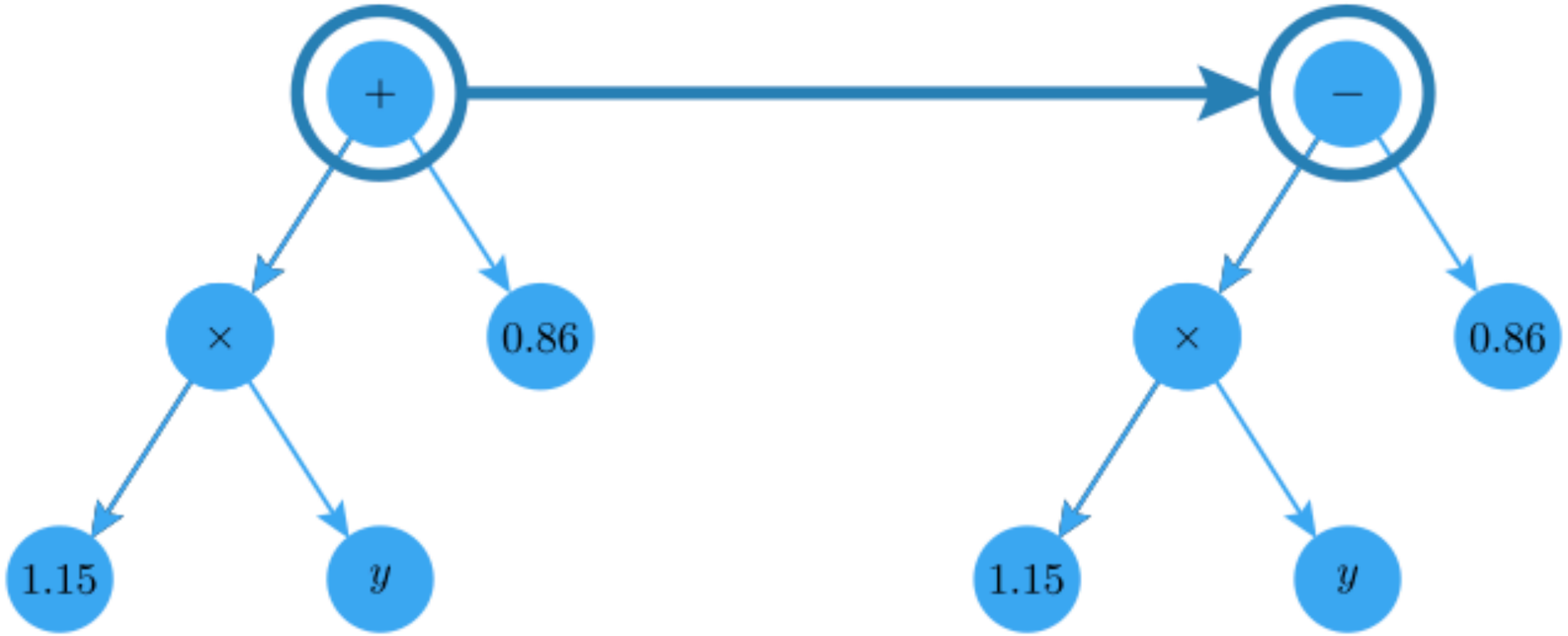
May 2, 2023

5 May 2023


Population P_k
selection
Parents
crossover
Offspring
mutation
Population P_{k+1}


x^x
$1.15y + 0.86$
Select Fittest
Mutation
or
(top two) Crossover
or
Simplification
or
Optimize
Constants
$1.15y^2 + 0.86$
or
$1.15y$
or ...
$1.15x^x + 0.86$
or
$1.15 + 0.86^x$
or ...
$1.15y + 0.86$
$1.1395y + 0.721$
Randomly
Subsample
x^x
$1.15y + 0.86$
xy^3
...
Replace
Oldest

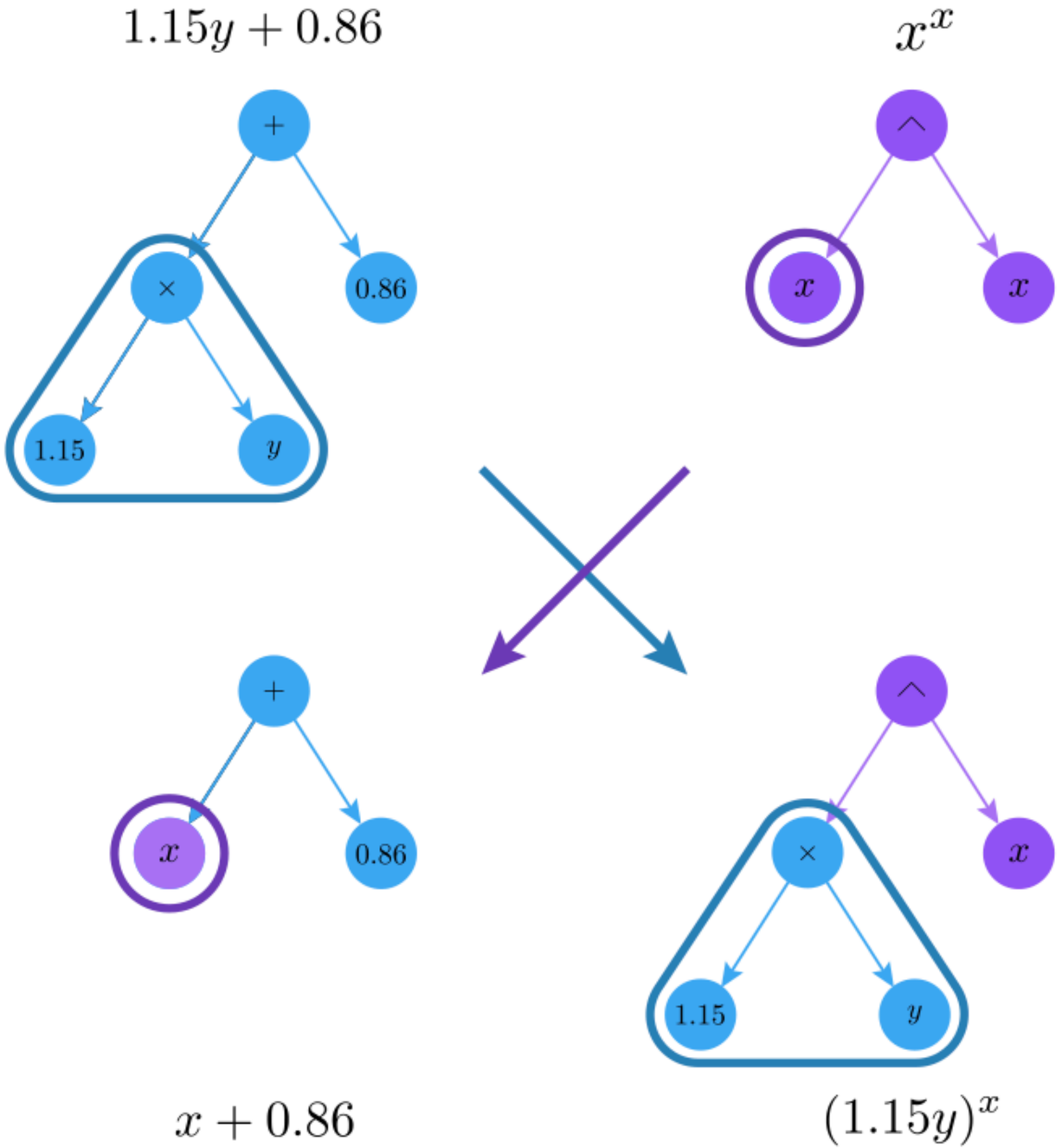
Mutation

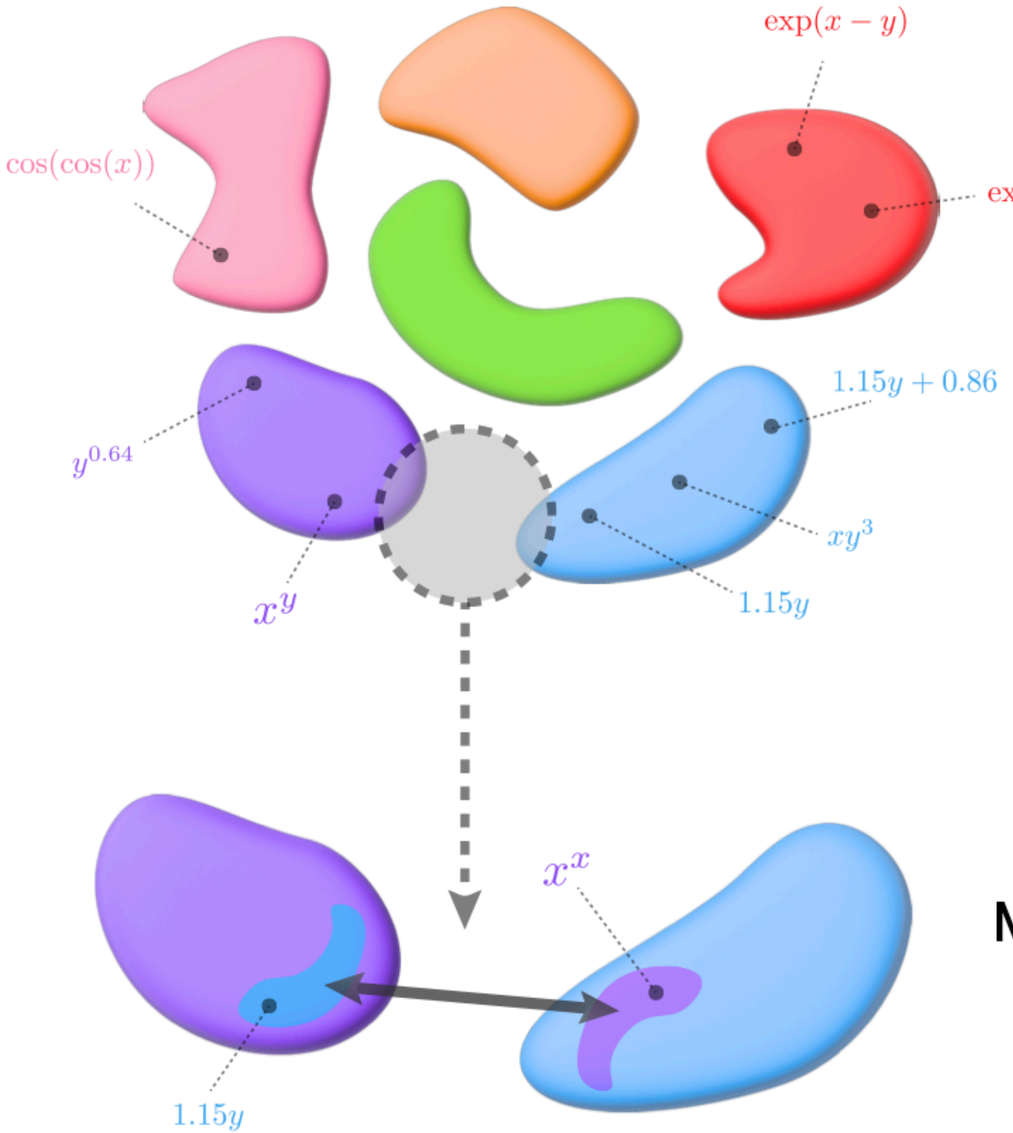


$1.15y + 0.86$

$1.15y - 0.86$

Crossover





Independent “islands” of expressions, each undergoing evolution

Migration between islands

Let's try it

dataset on course website

PySR