

Homework 1

Problem 1 [5pts]:

Suppose there are 4 independent variables. And 3 operations, i.e., addition, multiplication, and power, are allowed in connecting them into an expression. How many different expression can we have? For example, let the 4 variables be A, B, C, and D. Then $A+B * C ^D$ is one expression while $A^B * C+D$ is another. The order of variables matters, e.g., $A+B$ and $B+A$ are two different expressions. Each operator or variable appears in the expression EXACTLY ONCE. No parentheses.

The purpose of this problem is to understand how many different parametric equations can exist among a set of variables, and thus using traditional scientific discovery way to model the relationship between high-dimensional variables is very challenging.

Problem 2 [5pts (2.5pts for correct returns, and 2.5pts for correct plot)]:

In the example code for Unit 1, there is a demo that the score of a neural network changes along with the maximal number of iterations (i.e., the `max_iter` argument in the function `test_NN`). Now, let's visualize the change.

Implement a function `learning_curve` with the following I/O specifications:

```
def learning_curve(Ts, Hs, filename):
```

```
    # INSERT YOUR CODE HERE
```

```
    return max_iters, scores
```

where - the first two arguments `Ts` and `Hs` (1-D numpy array each, e.g., `[1,2,3]` not `[[1],[2], [3]]`) are the input and corresponding output for a supervised learning task, - the last argument `filename` (string) specifies the filename (Matplotlib uses the suffix to automatically determine the file format) to save the plot

How to submit

Just submit the modified `hw1.py` file. For problem 1, strictly just one line of comment at the top. Just the number. Then insert lines to finish function definition to `learning_curve` below. Feel free to import `numpy` and `matplotlib` in your function definition. Do NOT import modules beyond `numpy` and `matplotlib`.