DEEP LEARNING FOR PHYSICAL SYSTEMS ME504
ASSIGNMENT 3

Name: Shubha Tanaya
Entry no: 2021MEB1325

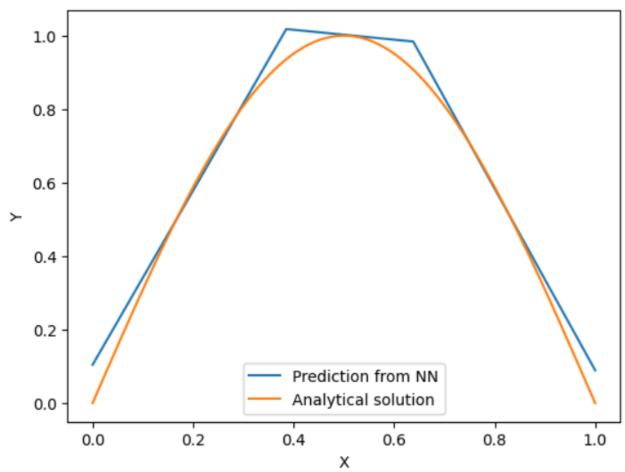
PART 1: REGRESSION PROBLEM

Function 1:
y=sin(πx)

Function 2:
y=2x+1

1. **Verify your regression program by plotting the predicted values against the analytical solution for both the functions.**

    y=sin(πx)
    1 Hidden layer with 8 nodes and relu activation function.

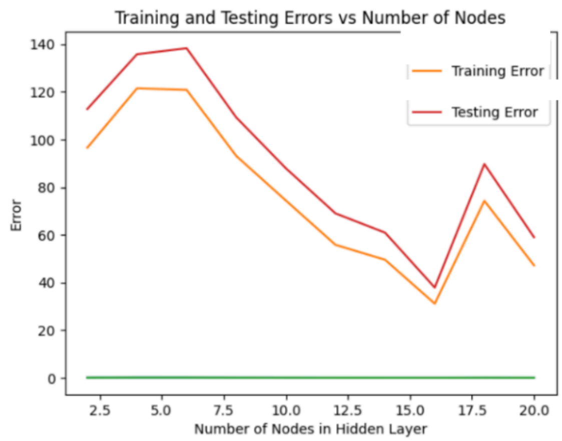


---

2. **Plot the training and testing error with respect to the number of nodes in layers. From this plot identify the optimal number of nodes. You can take any number of hidden layer for this case.**

y=sin(πx)

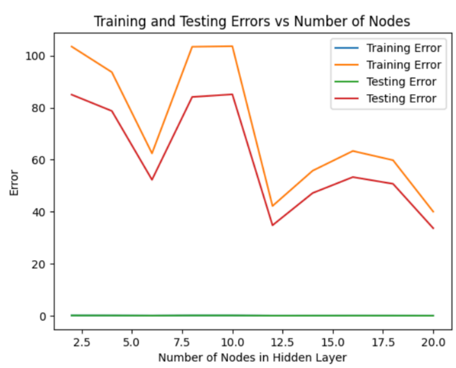
1 Hidden layer with [2, 4, 6, 8, 10, 12, 14, 16, 18, 20] nodes and relu activation function.



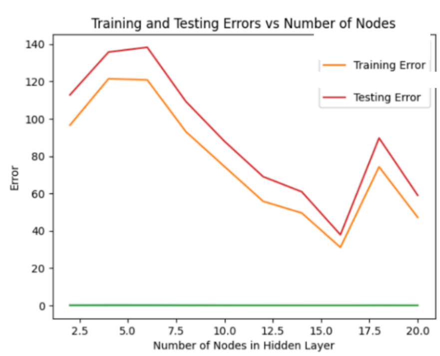
Optimal number of nodes according to the plot is 16.

---

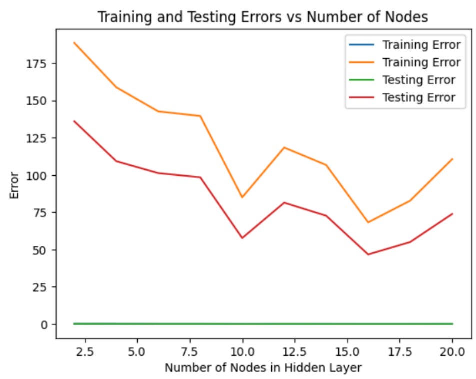
3. **Carry out the point no. 2 exercise for three different level of noise. Does the noise affects the conclusions of point-2.**



For noise 10%



For noise 5%



For noise 1%

- With higher noise, the optimal number of nodes increases to allow the model to capture the information effectively. However, there's a point where increasing the number of nodes further leads to overfitting rather than better generalization, when the model learns to capture the noise in the training data

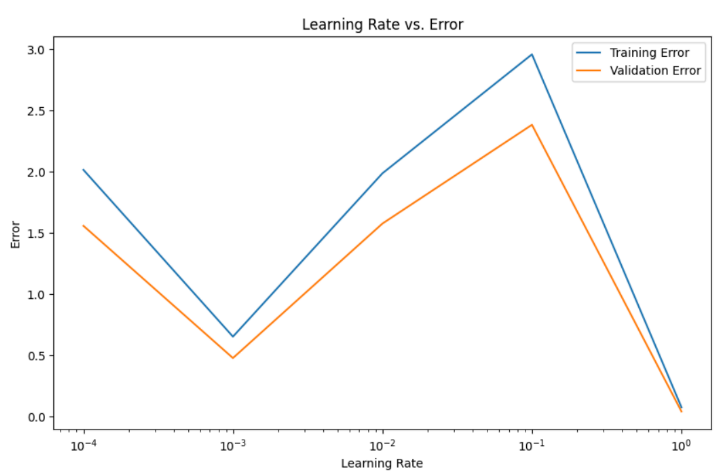
---

4. **Play with the learning rate and shows how it affects the overall performance. Does taking step learning rate is more helpful?**

- The learning rate determines the step size of weight updates during optimization while training. A small learning rate leads to slow convergence and a risk of local minima, while a large one may cause instability and overshooting. Finding the optimal learning rate is crucial so it's a very important hyperparameter.

For Function 2 variation of error with learning rate

y=2x+1

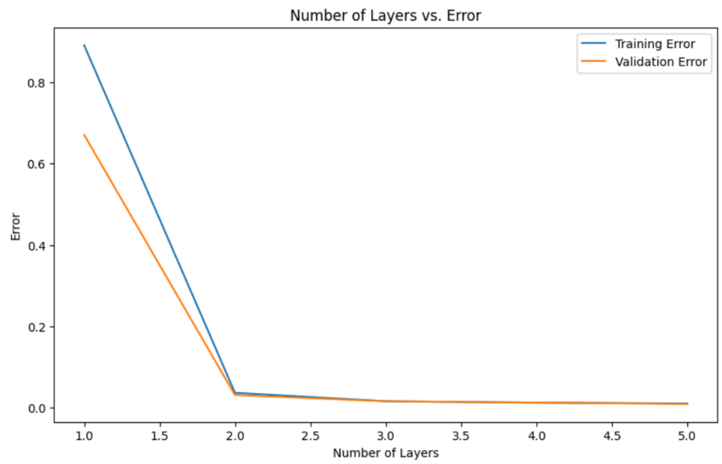


---

5. **Play with number of layers, and see its affect on the training and testing error.**

Increasing the number of layers enhances its capacity to learn the patterns of data, with more layers, the network can capture more features and information. Initially, adding layers reduces the training error, but too many layers can lead to overfitting, causing both training and testing errors to rise.

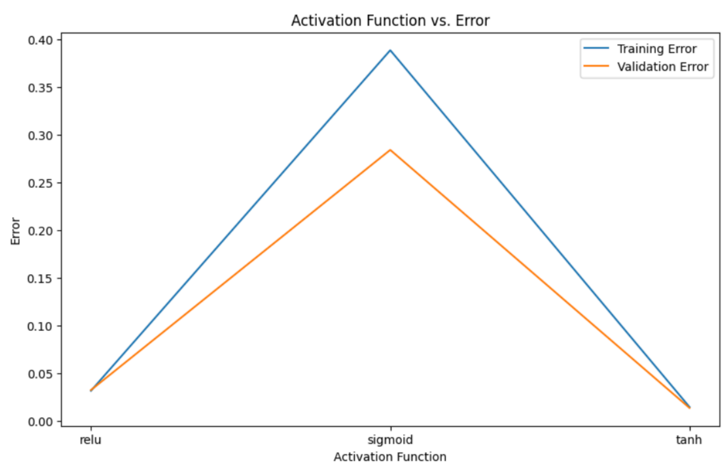
For Function 2:(y=2x+1) varying number of layers discretely by [1, 2, 3, 4, 5]



---

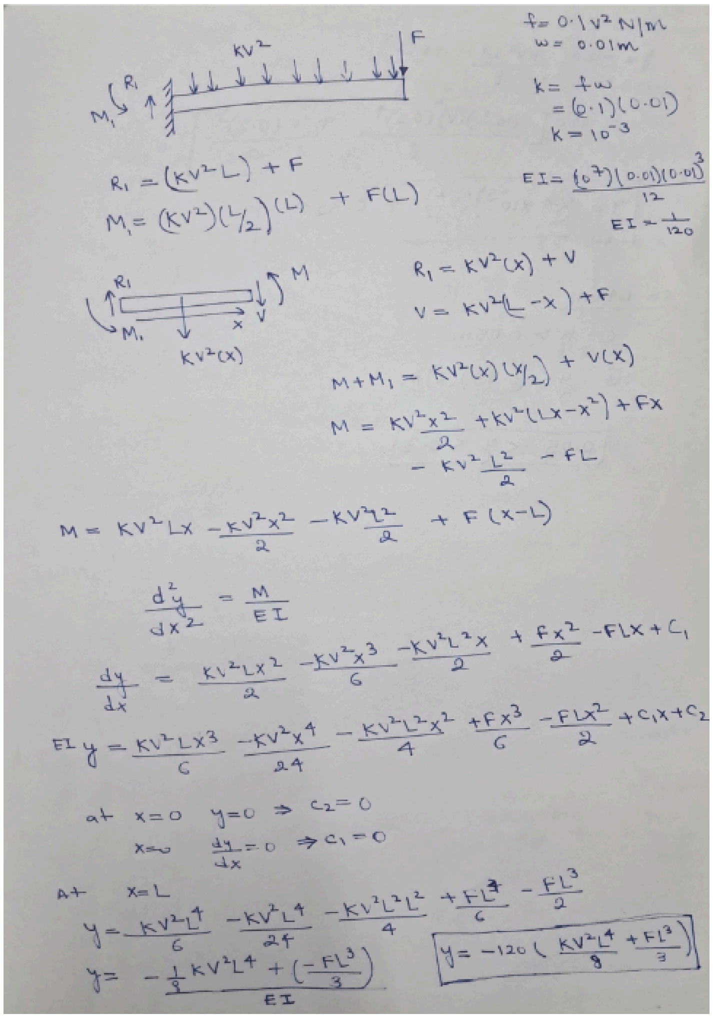
6. **Take different choices of activation functions, and analayze its effect.**

We will use 3 activation functions - 'relu', 'sigmoid', 'tanh' for function 2:(y=2x+1)



- ReLU and tanh activation functions result in lower training and validation errors compared to the sigmoid function. Sigmoid performs the worst, mainly due to the vanishing gradient problem.

PART 2: CLASSIFICATION PROBLEM:



$f = 0.1\,v^2$ N/m
$w = 0.01$ m

KV^2 F R_1 M_1

$k = fw$
$= (0.1)(0.01)$
$k = 10^{-3}$

$R_1 = (KV^2 L) + F$

$M_1 = (KV^2)(L/2)(L) + F(L)$

$EI = \frac{(10^7)(0.01)(0.01)^3}{12}$

$EI = \frac{1}{120}$

R_1 M M_1 x V $KV^2(x)$

$R_1 = KV^2(x) + V$

$V = KV^2(L - x) + F$

$M + M_1 = KV^2(x)(x/2) + V(x)$

$M = \frac{KV^2x^2}{2} + KV^2(Lx - x^2) + Fx - \frac{KV^2L^2}{2} - FL$

$M = KV^2Lx - \frac{KV^2x^2}{2} - \frac{KV^2L^2}{2} + F(x-L)$

$$\frac{d^2y}{dx^2} = \frac{M}{EI}$$

$$\frac{dy}{dx} = \frac{KV^2Lx^2}{2} - \frac{KV^2x^3}{6} - \frac{KV^2L^2x}{2} + \frac{Fx^2}{2} - FLx + C_1$$

$$EI\,y = \frac{KV^2Lx^3}{6} - \frac{KV^2x^4}{24} - \frac{KV^2L^2x^2}{4} + \frac{Fx^3}{6} - \frac{FLx^2}{2} + C_1x + C_2$$

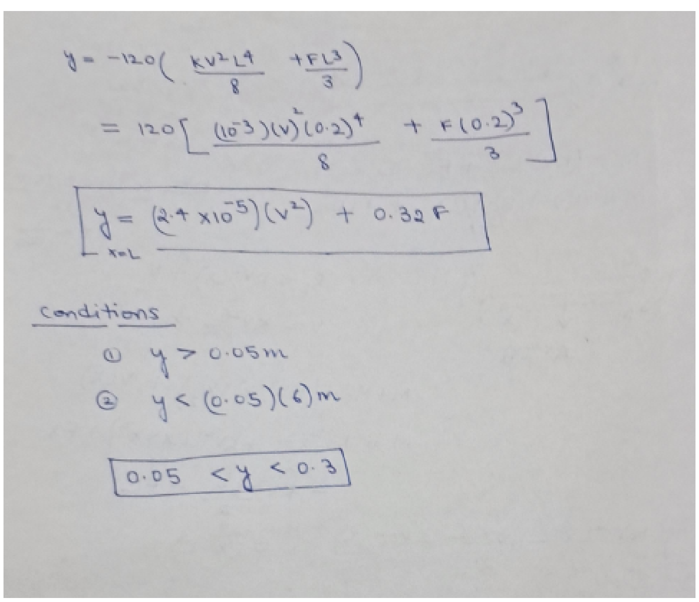
at $x=0$ $y=0 \Rightarrow C_2 = 0$
$x=0$ $\frac{dy}{dx}=0 \Rightarrow C_1 = 0$

At $x=L$

$$y = \frac{KV^2L^4}{6} - \frac{KV^2L^4}{24} - \frac{KV^2L^2L^2}{4} + \frac{FL^3}{6} - \frac{FL^3}{2}$$

$$y = \frac{-\frac{1}{8}KV^2L^4 + \left(\frac{-FL^3}{3}\right)}{EI}$$

$$\boxed{y = -120\left(\frac{KV^2L^4}{8} + \frac{FL^3}{3}\right)}$$



$$y = -120\left(\frac{KV^2L^4}{8} + \frac{FL^3}{3}\right)$$

$$= 120\left[\frac{(10^{-3})(v)^2(0.2)^4}{8} + \frac{F(0.2)^3}{3}\right]$$

$$\boxed{y\big|_{x=L} = (2.4\times10^{-5})(v^2) + 0.32F}$$

Conditions

① $y > 0.05$ m
② $y < (0.05)(6)$ m

$$\boxed{0.05 < y < 0.3}$$

Analytical Solution of the Physical System, model is formed similar to Part 1, just using labels this time.