Homework 1

David A. Halloran

January 15, 2022

Theory

1) Given the following calculate h:

$$x = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

$$W = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

$$b = \begin{bmatrix} -1 & 2 \end{bmatrix}$$

First multiply x into W which gives:

$$1 * 1 + 2 * 3 + 3 * 5 = 22$$

$$1 * 2 + 2 * 4 + 3 * 6 = 28$$

Simplified it looks like:

$$h = \begin{bmatrix} 22 & 28 \end{bmatrix} + b$$

Next we add the biases which gives us an answer of:

$$h = \begin{bmatrix} 21 & 30 \end{bmatrix}$$

2) Given matrix $h = \begin{bmatrix} 10 & -1 \end{bmatrix}$ calculate the output of feeding this into different activation functions:

2a) Linear:

$$g(h) = h = h = \begin{bmatrix} 10 & -1 \end{bmatrix}$$

2b) RELU: g(h) = max(0, h)

$$h[0] = 10$$

$$max(0, 10) = 10$$

$$h[1] = -1$$

$$max(0, -1) = 0$$

$$h = \begin{bmatrix} 10 & 0 \end{bmatrix}$$

2c) Sigmoid:

$$h[0] = \frac{1}{1 + e^{-10}} = 0.9999546021$$

$$h[1] = \frac{1}{1 + e^{1}} = 0.2689414214$$

$$h = \begin{bmatrix} 0.9999546021 & 0.2689414214 \end{bmatrix}$$

2d) Hyperbolic Tangent:

$$h[0] = \frac{e^{10} - e^{-10}}{e^{10} + e^{-10}} = 0.9999999959$$

$$h[1] = \frac{e^{-1} - e^{1}}{e^{-1} + e^{1}} = -0.761594156$$

$$h = \begin{bmatrix} 0.9999999959 & -0.761594156 \end{bmatrix}$$

2e) Softmax:

$$h[0] = \frac{e^{10}}{e^{10} + e^{-1}} = 0.9999832986$$

$$h[1] = \frac{e^{-1}}{e^{10} + e^{-1}} = 0.0000167014218$$

$$h = \begin{bmatrix} 0.9999832986 & 0.0000167014218 \end{bmatrix}$$

Testing The Layers

This section will include the outputs of different layers with the test matrix:

$$x = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$$

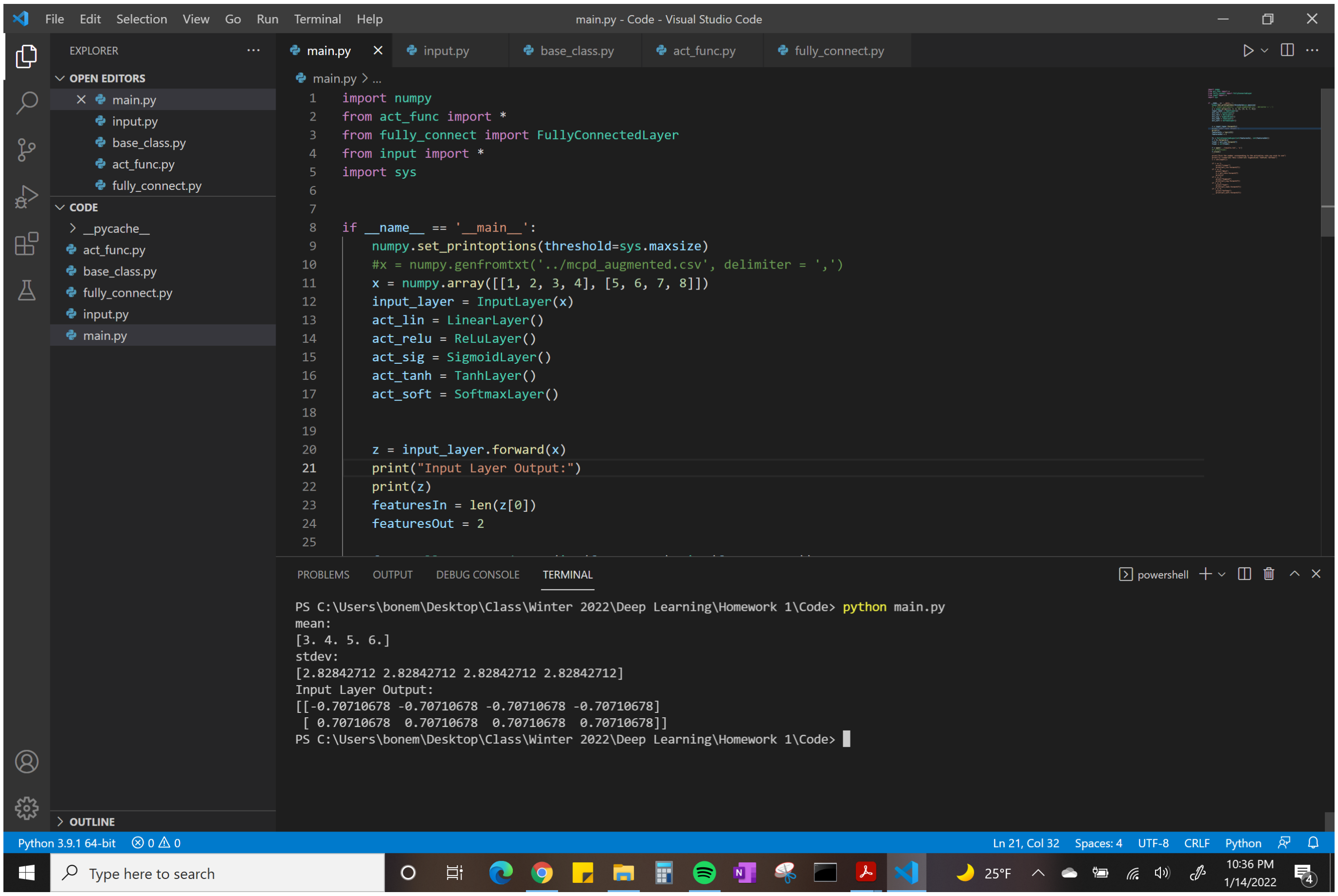

Figure 1: Input Layer

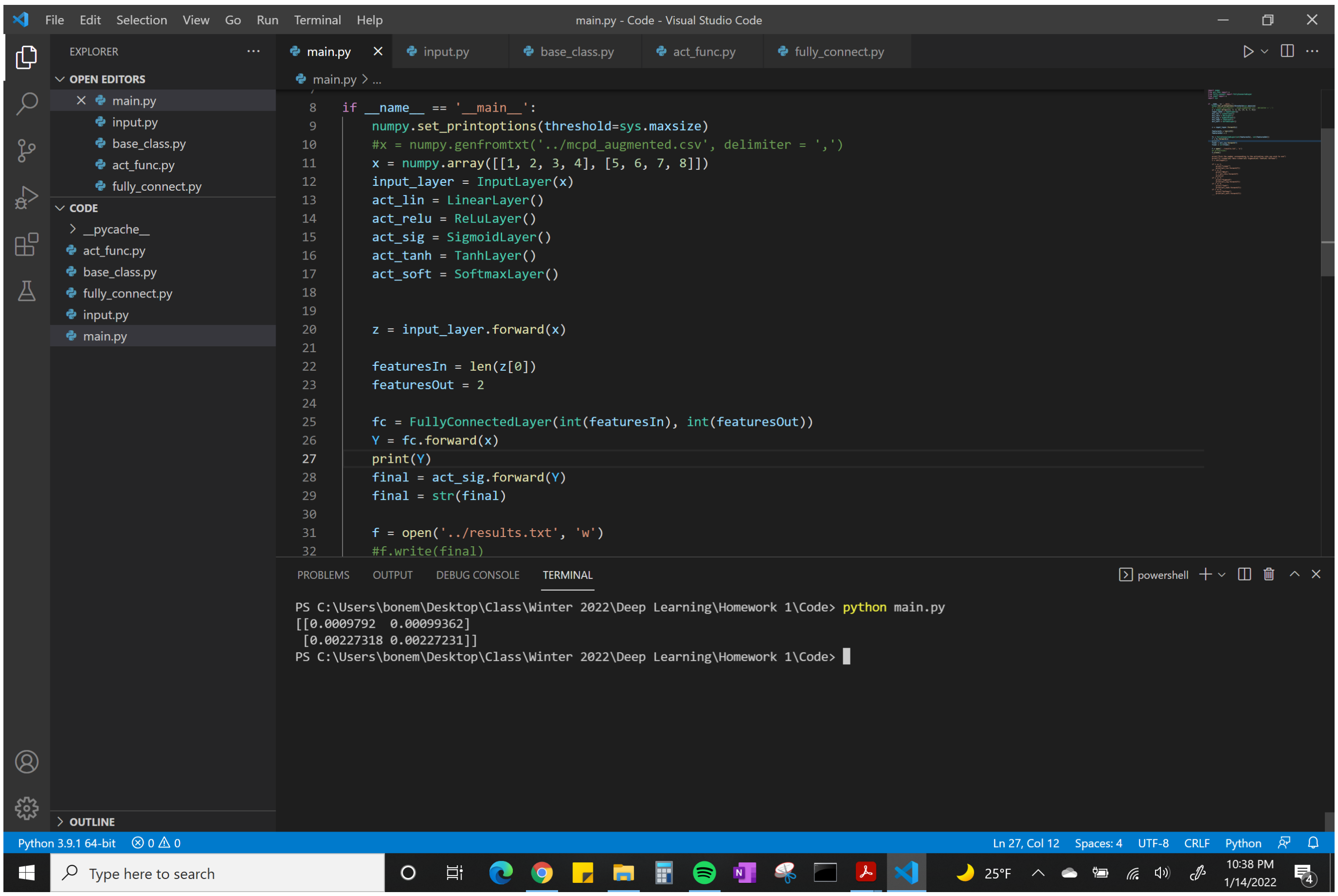

Figure 2: Fully Connected Layer

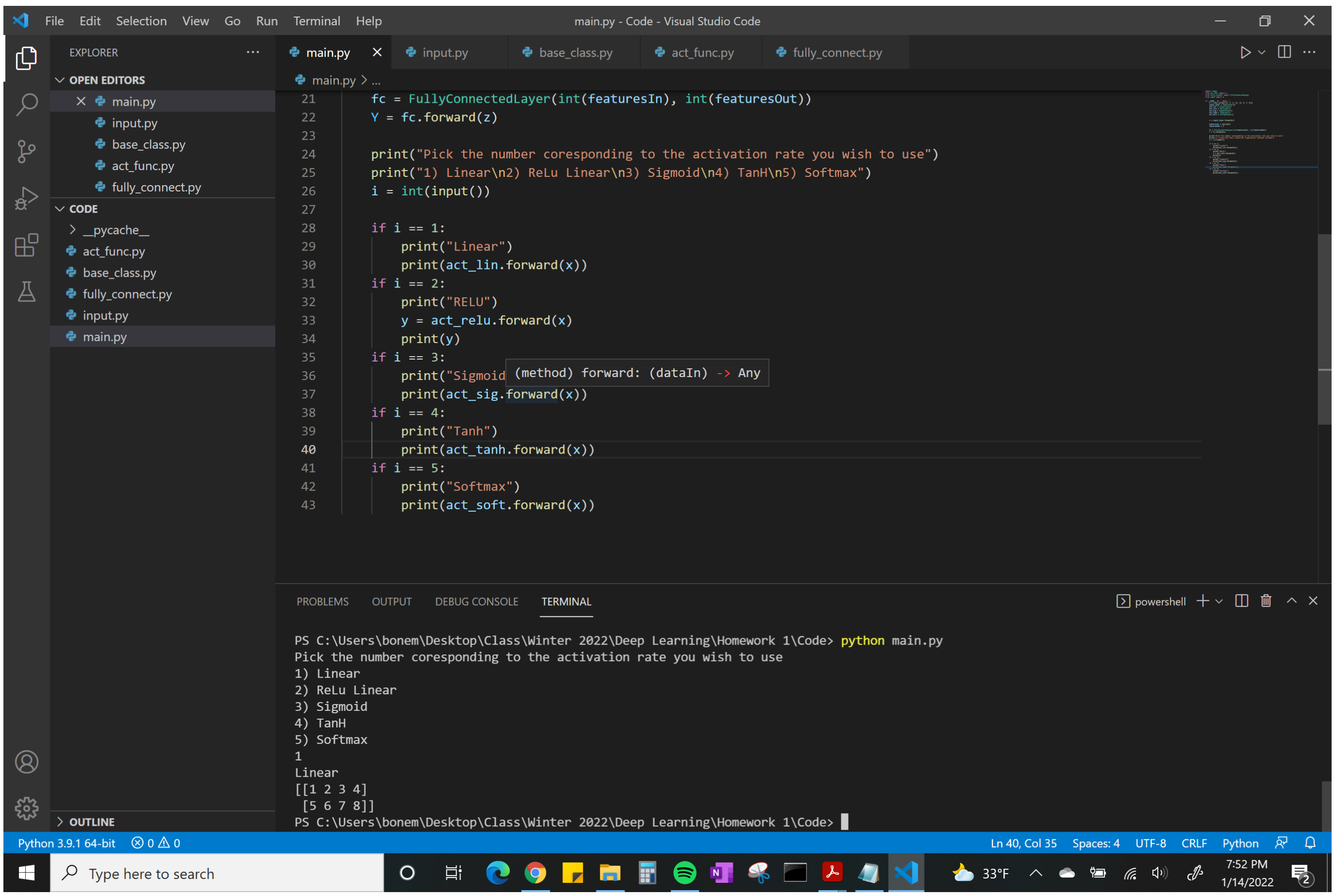

Figure 3: Linear Layer

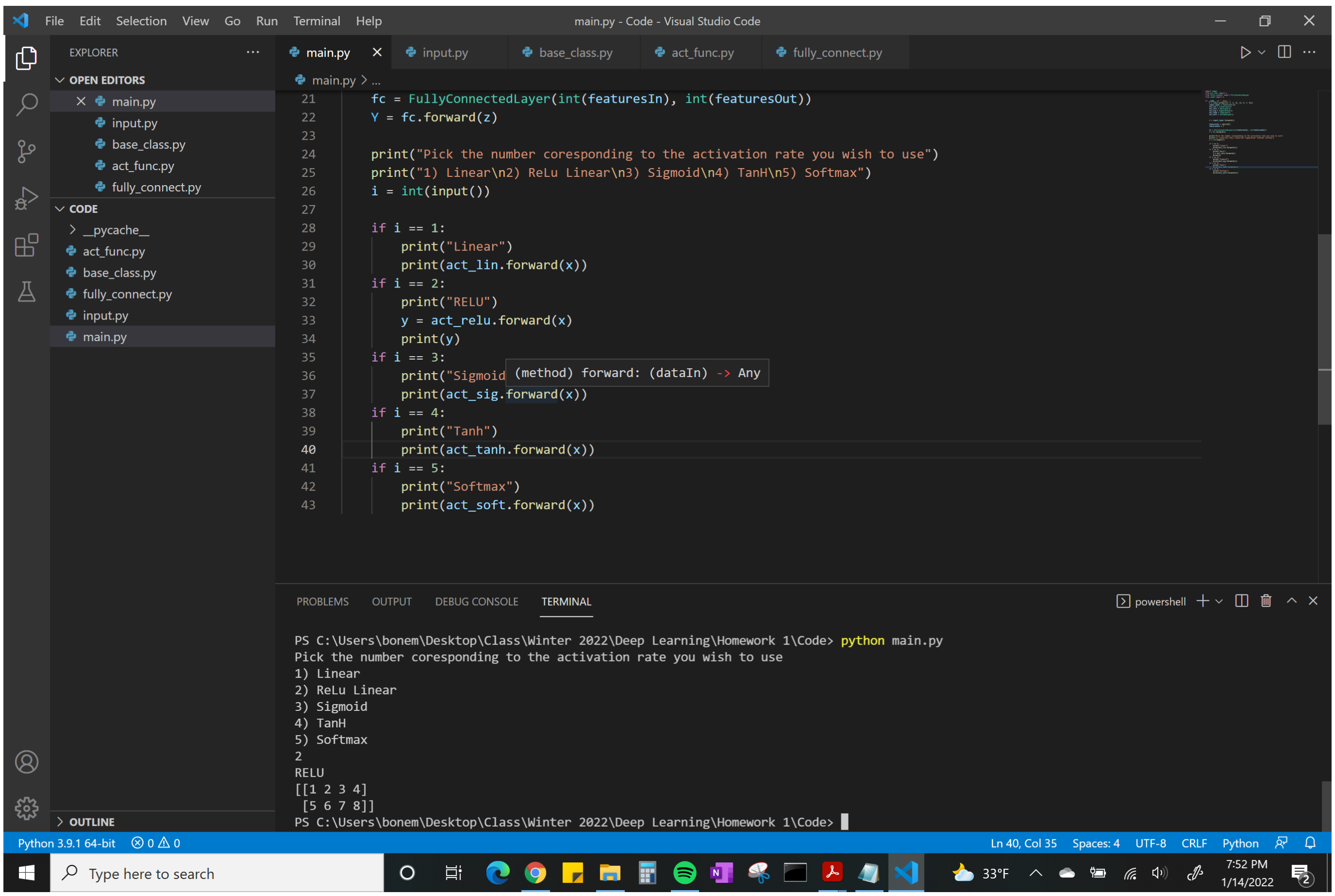

Figure 4: ReLu Layer

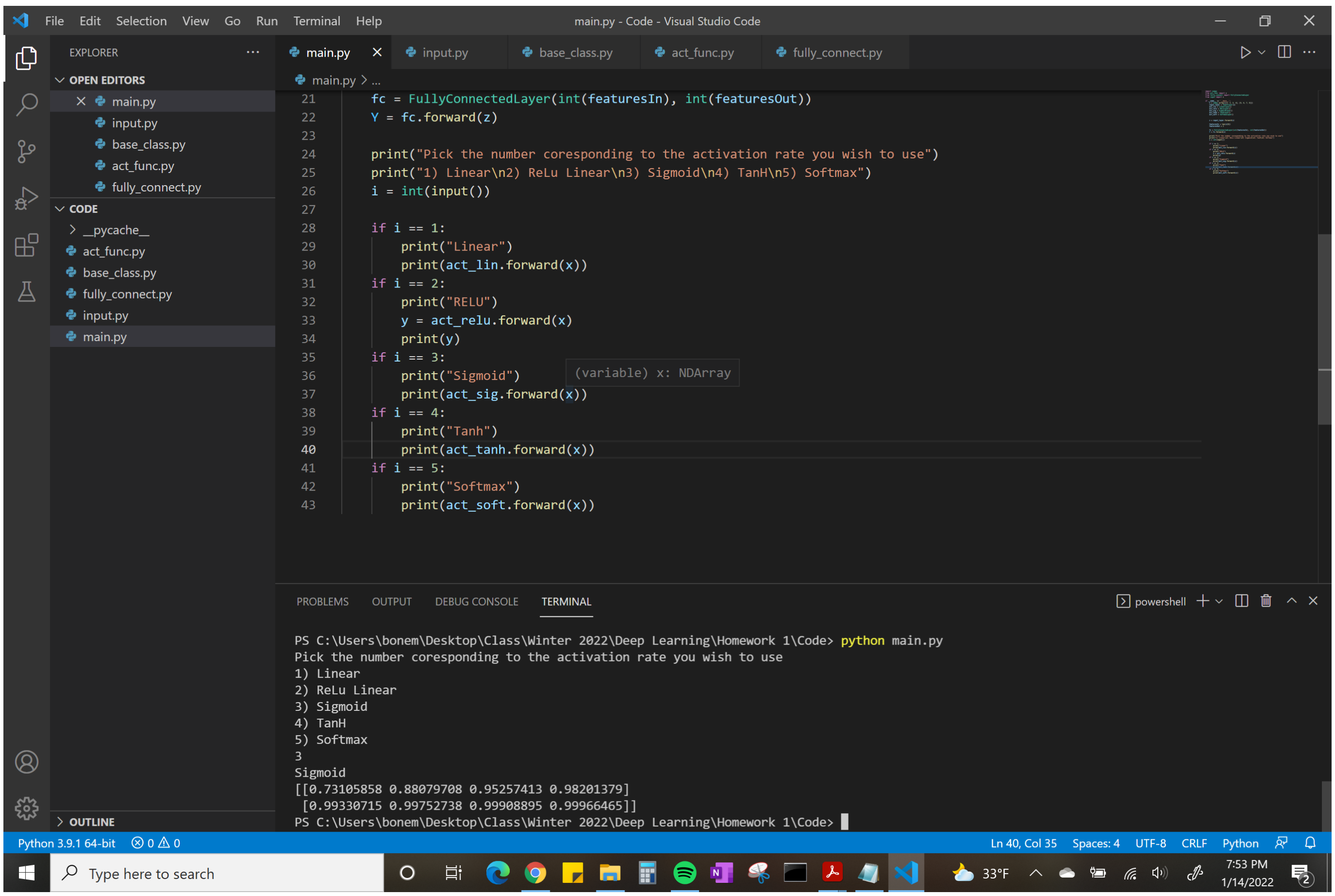

Figure 5: Sigmoid Layer

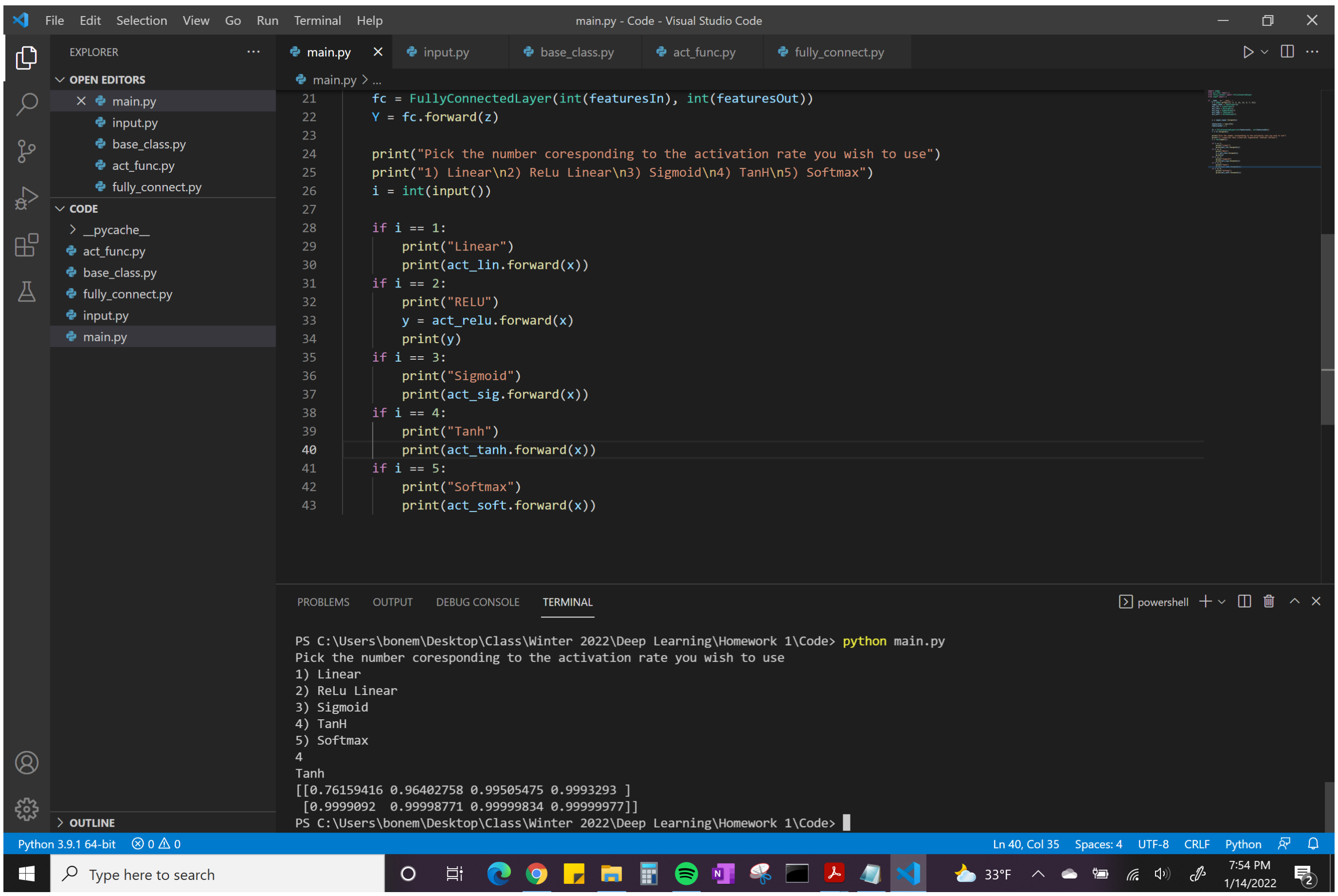

Figure 6: TanH Layer

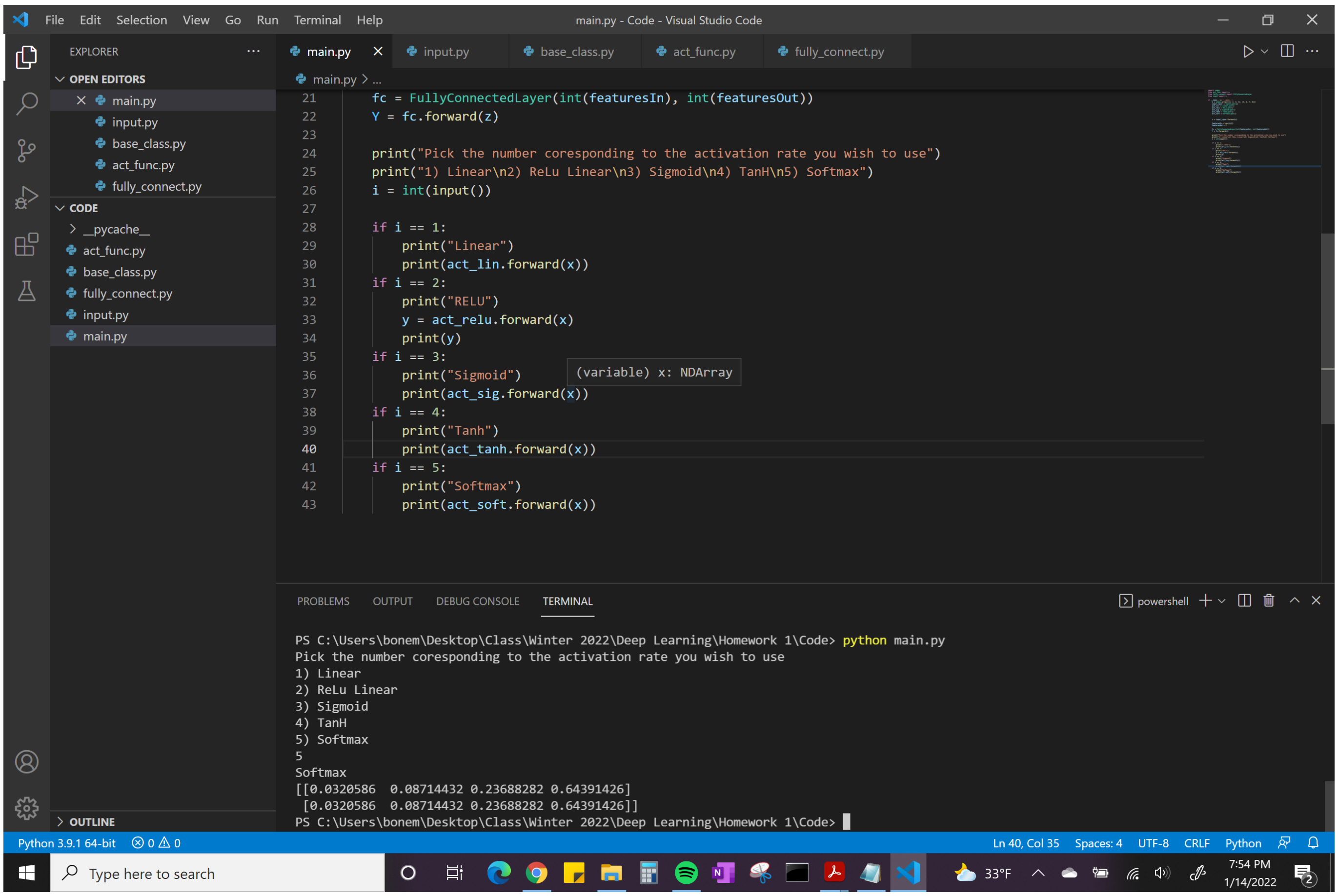

Figure 7: Softmax Layer

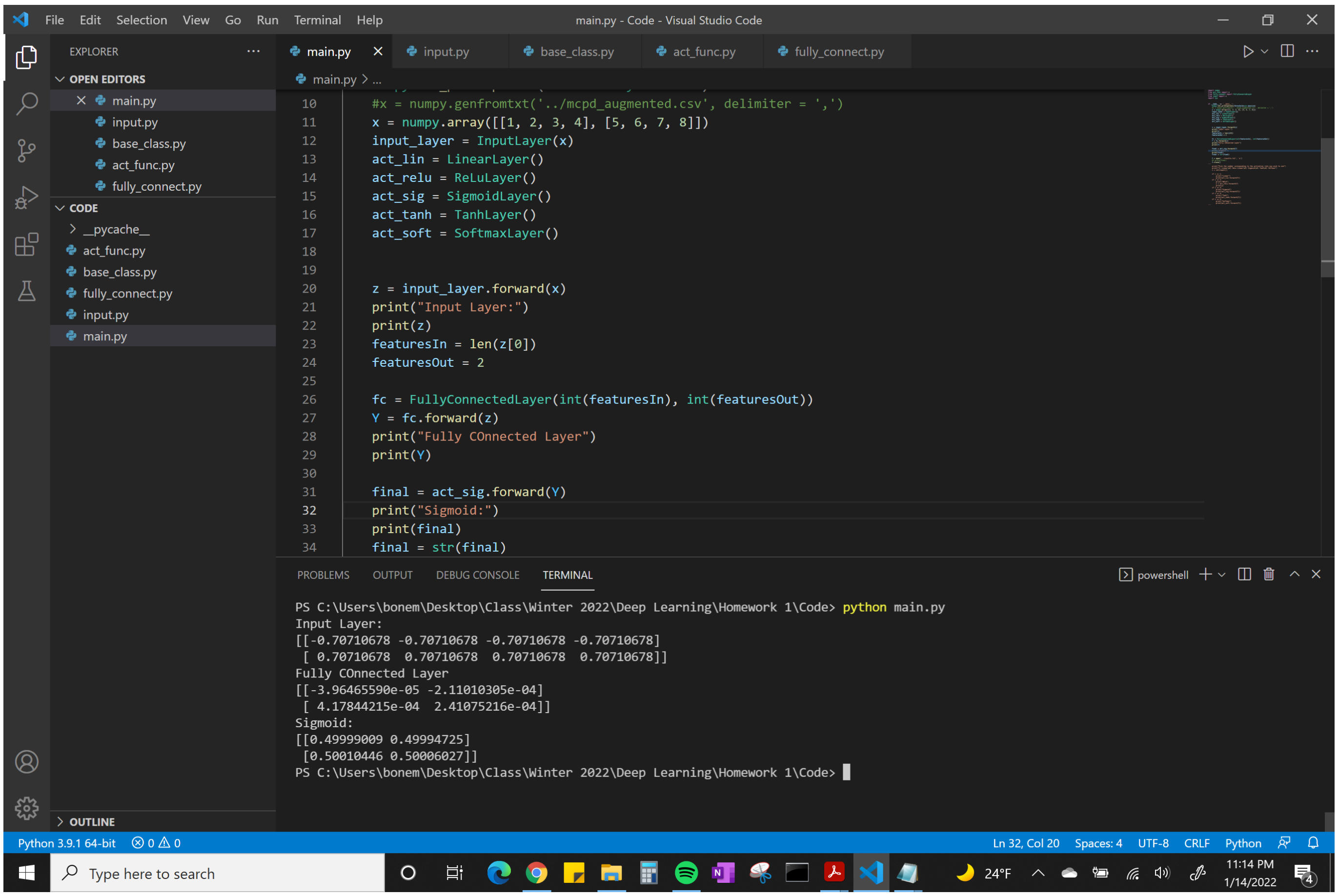

Figure 8: From Input to FC to Sigmoid

Medical Data

For the final test the program returns a 1338 x 2 matrix with the very first observation as the following:

$$h[0] = \begin{bmatrix} 0.4999867 & 0.50005346 \end{bmatrix}$$