CS 6375
ASSIGNMENT 2 Report

Names of students in your group:
Yash Nadkarni (ydn200000)
Soham Savalapurkar (sxs200389)

Number of free late days used: 0

Details about our Dataset:

For assignment 2, we used the banknote authentication dataset from the UCI website. This dataset has 4 independent variables and 1 dependent (target) variable. It has 1372 instances.

Independent Variables:

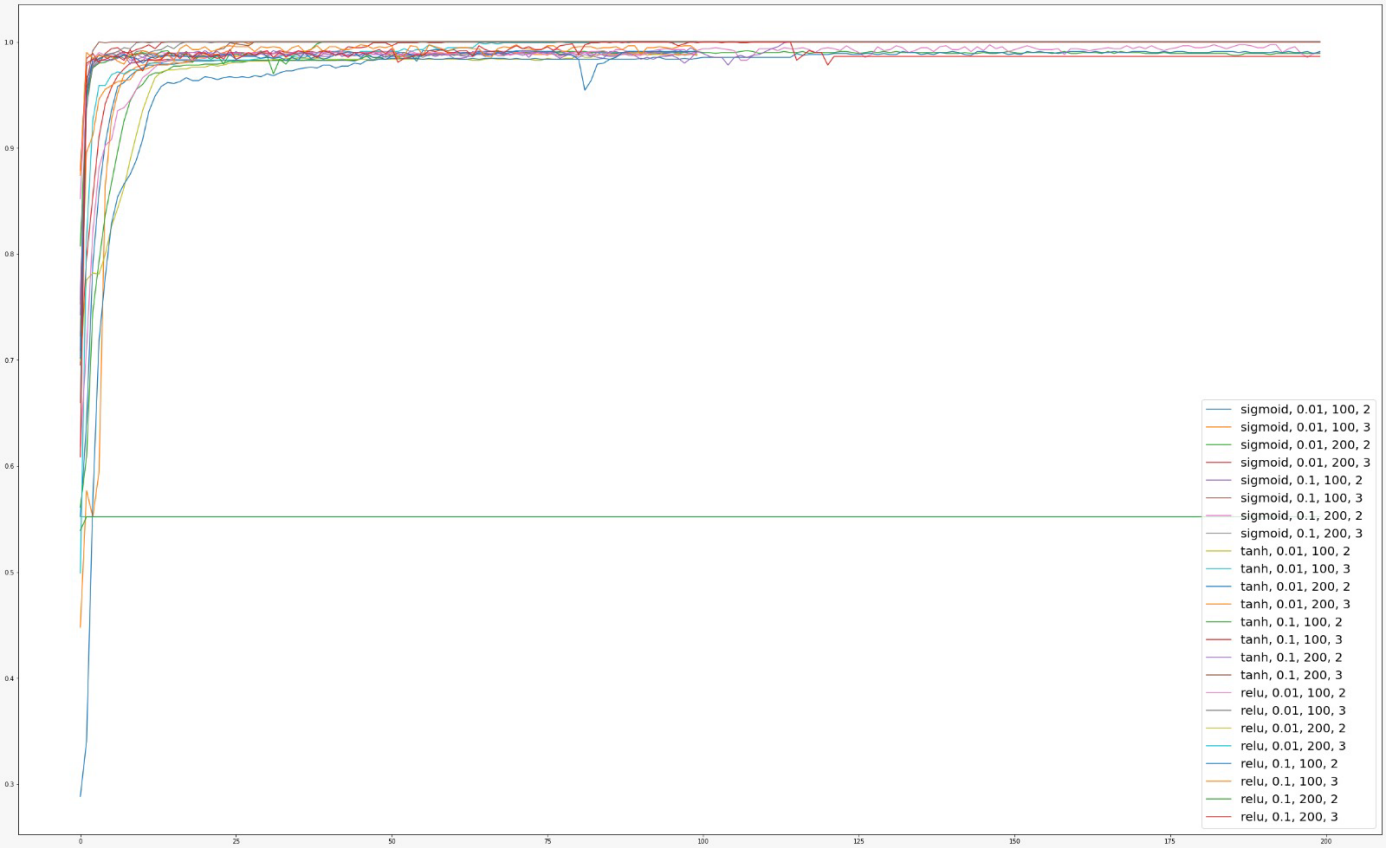
1. Variance of Wavelet
2. Skewness of Wavelet
3. Curtosis of Wavelet
4. Entropy of Image

Target Variable:

1. Class(integer)

For more information, check here.

Model History Plot of our model:



Output in tabular format:

1. Output for different combination of parameters

	activation	learning_rate	epochs	num_hidden_layers	training_accuracy	\
0	sigmoid	0.01	100	2	0.990876	
1	sigmoid	0.01	100	3	0.987226	
2	sigmoid	0.01	200	2	0.989051	
3	sigmoid	0.01	200	3	1.000000	
4	sigmoid	0.10	100	2	0.990876	
5	sigmoid	0.10	100	3	1.000000	
6	sigmoid	0.10	200	2	0.989964	
7	sigmoid	0.10	200	3	1.000000	
8	tanh	0.01	100	2	0.988139	
9	tanh	0.01	100	3	1.000000	
10	tanh	0.01	200	2	0.990876	
11	tanh	0.01	200	3	1.000000	
12	tanh	0.10	100	2	1.000000	
13	tanh	0.10	100	3	1.000000	
14	tanh	0.10	200	2	1.000000	
15	tanh	0.10	200	3	1.000000	
16	relu	0.01	100	2	0.989051	
17	relu	0.01	100	3	1.000000	
18	relu	0.01	200	2	0.552007	
19	relu	0.01	200	3	0.552007	
20	relu	0.10	100	2	0.990876	
21	relu	0.10	100	3	0.991788	
22	relu	0.10	200	2	0.552007	
23	relu	0.10	200	3	0.986314	

2. Table for the training and testing errors

	testing_accuracy	training_error	testing_error
0	0.992727	0.008365	0.007273
1	0.992727	0.007792	0.007273
2	0.992727	0.007093	0.007273
3	1.000000	0.000115	0.000000
4	0.996364	0.006052	0.003636
5	1.000000	0.000014	0.000000
6	0.992727	0.009058	0.007273
7	1.000000	0.000006	0.000000
8	0.985455	0.009645	0.014545
9	1.000000	0.001569	0.000000
10	0.992727	0.008961	0.007273
11	1.000000	0.000057	0.000000
12	1.000000	0.000510	0.000000
13	1.000000	0.000016	0.000000
14	1.000000	0.000179	0.000000
15	1.000000	0.000003	0.000000
16	0.992727	0.007161	0.007273
17	1.000000	0.000006	0.000000
18	0.567273	0.447993	0.432727
19	0.567273	0.447993	0.432727
20	0.989091	0.008127	0.010909
21	0.996364	0.006371	0.003636
22	0.567273	0.447993	0.432727
23	0.978182	0.014015	0.021818

Brief Report Summary:

From the observations, we observed that on changing only the hyperparameters and keeping the model definition the same the accuracy is changed, and we obtain different values in most cases. From the observed results of the model, tanh performed the best. Though generally relu performs the best, on further research, we found that in few cases the relu function might be facing the dying relu problem, due to which the accuracy for some cases is low.

Assumptions: None

References:

https://stats.stackexchange.com/questions/126238/what-are-the-advantages-of-relu-over-sigmoid-function-in-deep-neural-networks

Referred the following link for dying relu problem.