CS 4375
ASSIGNMENT 2. Report

Names of students in your group:

Haeun Kim

Number of free late days used: 2

Note: You are allowed a **total** of 4 free late days for the **entire semester**. You can use at most 2 for each assignment. After that, there will be a penalty of 10% for each late day.

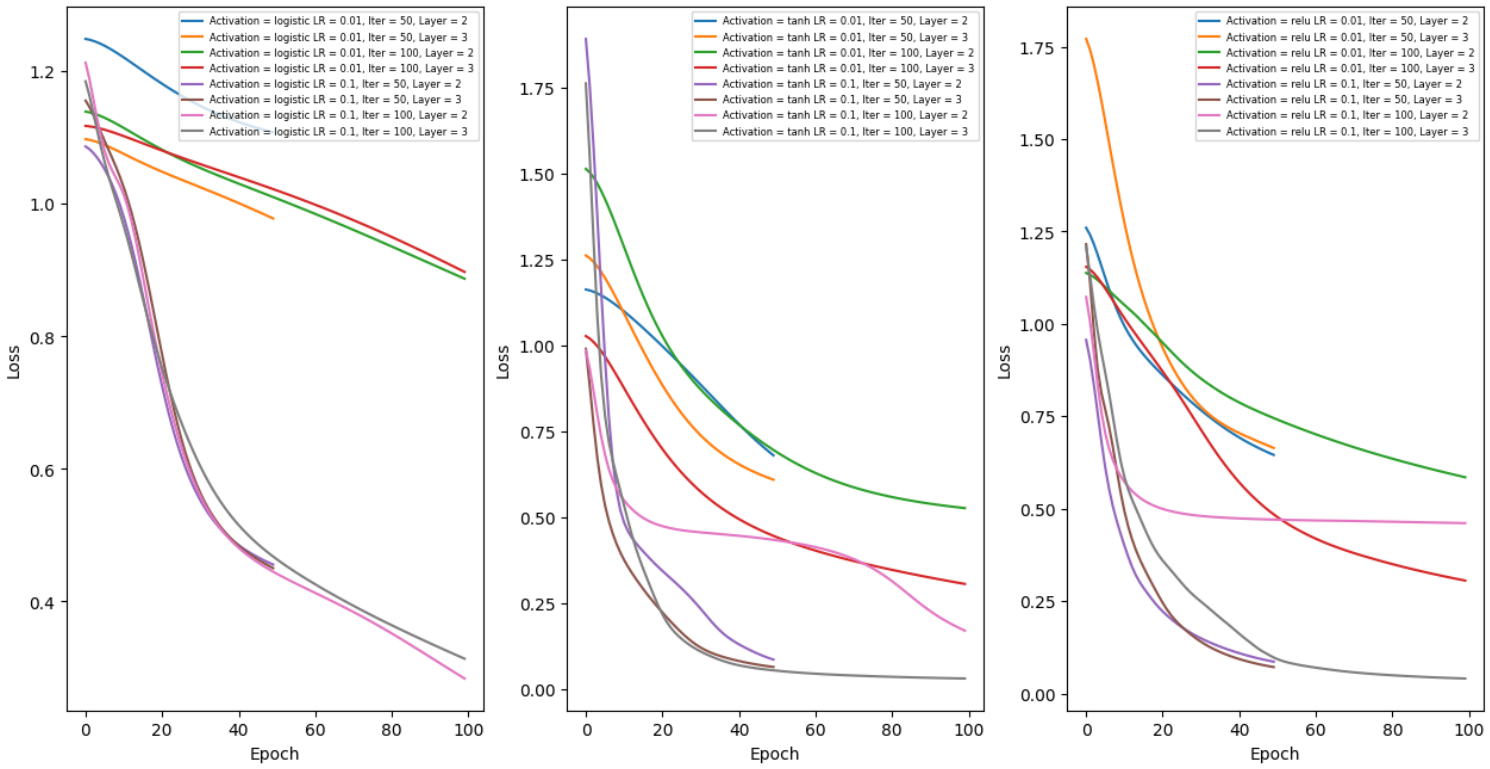
Please list clearly all the sources/references that you have used in this assignment.

Dataset:

- https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data
- Course resourse ("Scikit-Learn Lab")
- https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html
- https://stackoverflow.com/questions/39770376/scikit-learn-get-accuracy-scores-for-each-class

Report Part.2

I did this homework with IRIS dataset from UCI machine learning repository. It has 150 Instances, and 5 attributes(sepal length, sepal width, petal, length, petal width, and class). In the preprocess part, I renamed the columns and deleted NA or duplicates. The speices which I do classification is categorical, so I changed it into numerical by using the LabelEncoder from sklearn.preprocessing. I also did the standarization by using the StandardScaler from sklearn.preprocessing. I split the dataset 80(train) : 20(test).



I worked with all possible combination of hyperparameters(activation, iteration, learning rate, and hidden layers) by using MLPClassifier from sklearn.neural_network. Above plot is the result of track of model history. I used 'loss_curve_' from the MLPClassifier, and I adjust the iteration into [50, 100] because some accuracy plot was cut in some points when I worked with iteration [100, 200]. This plot is Loss vs Epoch. Loss is '1-accuracy', so I could think that 'Activation = logistic, learning rate = 0.1, iteration = 100, and hidden layer = 2', 'Activation = tanh, learning rate = 0.1, iteration = 100, and hidden layer = 3', and 'Activation = relu, learning rate = 0.1, iteration = 100, and hidden layer = 3' have the high accuracy.

	activation	learn_rate	epochs	num_hidden_layer	acc_train	loss_train	MSE_train	acc_test	loss_test	MSE_test
0	logistic	0.01	50	2	0.589744	1.032775	0.641026	0.566667	0.433333	0.633333
1	logistic	0.01	50	3	0.666667	0.894570	0.333333	0.633333	0.366667	0.366667
2	logistic	0.01	100	2	0.726496	0.866461	0.273504	0.766667	0.233333	0.233333
3	logistic	0.01	100	3	0.581197	0.973287	0.418803	0.666667	0.333333	0.333333
4	logistic	0.10	50	2	0.683761	0.697634	0.316239	0.700000	0.300000	0.300000
5	logistic	0.10	50	3	0.914530	0.397968	0.085470	0.966667	0.033333	0.033333
6	logistic	0.10	100	2	0.897436	0.410937	0.102564	0.933333	0.066667	0.066667
7	logistic	0.10	100	3	0.897436	0.378265	0.102564	0.933333	0.066667	0.066667
8	tanh	0.01	50	2	0.683761	0.707934	0.341880	0.633333	0.366667	0.366667
9	tanh	0.01	50	3	0.820513	0.507191	0.179487	0.866667	0.133333	0.133333
10	tanh	0.01	100	2	0.829060	0.436226	0.170940	0.900000	0.100000	0.100000
11	tanh	0.01	100	3	0.829060	0.413567	0.170940	0.866667	0.133333	0.133333
12	tanh	0.10	50	2	0.982906	0.143756	0.017094	0.966667	0.033333	0.033333
13	tanh	0.10	50	3	0.974359	0.117173	0.025641	0.966667	0.033333	0.033333
14	tanh	0.10	100	2	0.982906	0.101551	0.017094	0.966667	0.033333	0.033333
15	tanh	0.10	100	3	0.991453	0.049761	0.008547	0.966667	0.033333	0.033333
16	relu	0.01	50	2	0.794872	0.718176	0.205128	0.766667	0.233333	0.233333
17	relu	0.01	50	3	0.914530	0.400094	0.085470	0.933333	0.066667	0.066667
18	relu	0.01	100	2	0.863248	0.569457	0.136752	0.933333	0.066667	0.066667
19	relu	0.01	100	3	0.666667	0.528202	0.333333	0.666667	0.333333	0.333333
20	relu	0.10	50	2	0.991453	0.107352	0.008547	0.966667	0.033333	0.033333
21	relu	0.10	50	3	0.974359	0.096689	0.025641	0.966667	0.033333	0.033333
22	relu	0.10	100	2	0.991453	0.069377	0.008547	0.966667	0.033333	0.033333
23	relu	0.10	100	3	0.991453	0.047340	0.008547	0.966667	0.033333	0.033333

I think Relu is the best activation. This is because, when comparing accuracy, loss, and MSE with other, the Relu's accuracy is higher, loss is lower, and MSE is lower than others. The best hyperparameter within the activation is that: For logistic [learning rate 0.10, iteration 50, hidden layer 3], For tanh [learning rate 0.10, iteration 100, hidden layer 3], For relu [learning rate 0.10, iteration 100, hidden layer 3]. In this Iris dataset, when I do the neural network learning, I could conclude that the optimized activation is Relu, the optimized learning rate is 0.10, and the optimized hidden layer is 3.