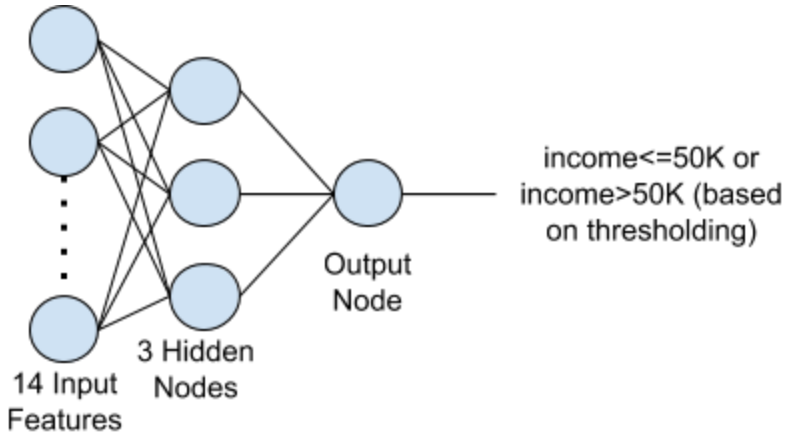
Statistical Machine Learning **Assignment 4** **Due Date: April 9, 2018**

Theory questions:
Q1: Question 4 from section 3.2 of textbook.
Q2: Question 10 from section 3.2 of textbook.

Programming questions:
Q1: Implement the following Neural Networks (from scratch with backpropagation):



Dataset

- UCI Adult Dataset. Link: https://archive.ics.uci.edu/ml/datasets/Adult
- The task is to predict if a person earns over 50k a year or not. The dataset has 14 attributes. Use these 14 attributes to predict the income range.

Architecture and protocol

1. Hidden node activation: ReLu, output node activation: Linear
2. Number of nodes in the hidden layer: 3
3. Task: Perform 2 class classification to predict if a person earns over 50k a year or not. For this task, you may binarize the output obtained from the output layer into 0 or 1 using a threshold.
4. Based on the output of the system after thresholding, you have to back propagate the error and update the weights of the connections between the nodes.
5. Use 50% samples of each class for training (odd samples) and rest 50% for testing (even samples).

Deliverables

1. A running code with the above defined architecture.
2. A confusion matrix to illustrate correct and incorrect classifications for all testing samples. Also report class wise and overall accuracy.
3. ROC plot between two classes (income<=50k and income>50k).
4. Report equal error rate (EER) on the testing set.

Q2: EMNIST (Extended MNIST Balanced dataset) classification into 47 classes using 3 hidden layer NN. The base architecture should be: [#input, H1: 256 nodes, H2: 128 nodes, H3: 64 nodes, Output: 47 nodes]. You may use any inbuilt library for NN.
Unless explicitly specified in the following parts, use: (a) linear activation, (b) learning rate=0.1, (c) epochs=100, and (d) # nodes as mentioned above.

1. Vary the learning rate 0.2, 0.1, 0.001. Analyze and plot the graph of accuracy vs the learning rate.
2. Vary number of nodes in the hidden layers two times. Analyze accuracies of base network vs increased network.
3. Add one more layer H4 of 128 nodes (after H3). Analyze accuracies of base network vs increased network.
4. Vary epochs from 20 to 100 with step size of 40. Analyze and plot the graph on accuracy vs epochs.
5. Perform the experiments with sigmoid and ReLu activation functions. Analyze and plot the graph on accuracy vs activation functions.
6. **BONUS:** Come Up with interesting experiments and analysis on the neural network which has not been asked in 1 to 5 tasks

Q3: Using CIFAR 10 dataset, implement the following using a simple one hidden layer Neural Network as a weak classifier:

Architecture of Neural Network [#input: 32*32 , H1: 50 nodes, Output: 10 nodes (number of classes in CIFAR 10)].

a) Implement boosting for classification.
b) Using the same framework implement bagging.
c) Use 5 and 15 such weak classifiers and report classification accuracies for both (a) and (b)
d) Analyze the results obtained in (a), (b), and (c).
e) Compare the accuracies obtained in (a) and (b) with a Multilayer Neural Network (2 hidden layers). Architecture of Neural Network [#input: 32*32 , H1: 400 nodes, H2: 150 nodes, Output: 10 nodes (number of classes in CIFAR 10)]. Analyze your results.

Link for CIFAR 10 : https://www.cs.toronto.edu/~kriz/cifar.html
Note: You can convert the images of CIFAR 10 into grayscale.

Deliverables: Codes for (a), (b) and (c). Write all the analysis and the results in your report. Results should be given in the report in proper tabular format.

Submission format:

1. All the graphs and figures should be labeled and placed in your PDF report. Use tables to improve readability of your report. The assignment would also include marks for the report and the comparative analysis of the results.
2. All the experimental results should be there in your report in proper tabular format. Report should be in a .pdf file.
3. Theory questions should be submitted in hard copy only in the drop box which will be placed a few days prior to the deadline.