AI Principles and Applications

Fall 2020 ------ Homework #4

Due Date: 9PM, Nov. 24th, 2020

Note: Submit all your answers included within a single .pdf file. Python code is to be used for listing only and it will not be executed from the submitted file.

Q1. Consider a guide and description for using SciKit Learn utilities ([kdd-site](https://www.kdnuggets.com/2016/10/beginners-guide-neural-networks-python-scikit-learn.html/2) and [help-file](https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html)) for creating neural network models from data. Also, consider the same data set that you used for Homework#3, described on a 2-D grid, given in the attached HW3Data.xlsx file. For each 2-D coordinate pair, the third column states the class to which this point belongs. The only possible classes are “0” and “1”. Perform the following tasks in the context of this neural network (NN) learning tool for creating models using the data in HW3Data.xls. For each part below submit the items shown in red font, **along with the Python code used for obtaining each result.**

1. (15) Select randomly 70% of the data points from HW3Data.xlsx. These points are to be used for learning the neural network. Learn the NN model for this training data, using only 2 neurons in the single hidden layer, and using default values for all other parameters. Use the learned model to predict the class for each of the 70 training data points. On a 2-D grid plot these training data points, showing the points for each predicted-class in a different color. On the plot, label/mark those data points whose predicted class is different from their actual/true class. How many data points have been misclassified?
2. (15) Use the remaining 30% of the data points in HW3Data.xlsx for testing. Test them with the NN model to find the predicted class label for each test data point. Use the information of the actual and the predicted class labels for each data point to create a confusion matrix for your learned model. (There are automated tools for creating the confusion matrix). Show the confusion matrix and the following performance metrics derived for the NN model: accuracy of the model, precision and recall for class “1”, and precision and recall for class “0”.
3. (20) Now consider all the points (natural numbers only) of the (70 X 70) 2-D grid as test points for the NN model. Find the predicted class label for each of the 4900 data points. Use programming help to generate the 4900 points and do not hand-code them. Create and show a display of the 70 X 70 2-D grid in which each point is displayed in color1 (say red) if its predicted class is “1”, and is displayed in color2 (say black) if its predicted class is “0”.
4. (10+10+15+15) Now repeat the steps 1 through 3 above with the difference (in step 1) that now you use 6 neurons in the single hidden layer instead of 2 neurons. Compare and contrast the boundaries obtained in the step #3 for the two cases of the number of neurons in the hidden layer. Also give your comments on the comparison of the class boundaries learned for decision tree in Homework#3 and the NN model boundaries learned in this assignment.