**CPE383 Machine Learning: Quiz8**

Use the “PCA Examples.ipynb” for the following:

1. 15 points. 1 hour. Using the Wine dataset, use PCA to reduce it to 2 PCA features.

1.1. Visualize the first 2 columns as a scatter plot using all (100%) of the data.

Chart, scatter chart

Description automatically generated

1.2. In another graph visualize the scatter plot for the 2 PCA features also using all the data.

Chart, scatter chart

Description automatically generated

1.3. What % of information (“explained variance”) is preserved in the 2 PCA features?

Graphical user interface, text

Description automatically generated

2. 10 points. 1 hour. Use 25% for testing for the Wine dataset. Plot a graph of the accuracy of the samples when using 1, 2, 3, ..., 13 PCA features along with showing how much % variance is preserved (p%) for each? Use logistic regression.

Graphical user interface, text, application

Description automatically generated

3. 15 points. 2 hours. Decision boundary. Plot the decision boundary for PCA-reduced 2 dimensions of the wine dataset. Use ML Perceptron in https://python-course.eu/machine-learning/neural-networks-with-scikit.php with 2 hidden layers.

Chart

Description automatically generated with medium confidence

4. Using the MNIST data of 784 dimensions (28 x 28 pixels).

4.1. 5 points. 0.5 hours. Show Quiz 7 problem #3 here again as a benchmark. Using 50,000 data samples for training and 20,000 for testing, apply standard scalar transform from the training dataset to both your training and testing datasets. Report the accuracy and print out the confusion matrix.

Graphical user interface, text, application

Description automatically generated

4.2. 20 points. 2 hours. Using the same neural network model as in 4.1 above, but using the PCA features that preserve up to 95% of the information content, find a new training and testing dataset using PCA transformation based on the training dataset. Apply standard scalar transform from the training dataset to both your training and testing datasets. Report the accuracy and print out the confusion matrix for PCA with reduced features. Graphical user interface, text

Description automatically generated