**MLPR - Lab 8**

INSTRUCTIONS

**Step 1:** Import all libraries

* Numpy
* Pandas
* Pyplot
* Self Training classifier from sklearn
* K Nearest neighbour classifier
* Confusion matrix, classification report, accuracy score, balanced accuracy score

**Step 2:** Read the data given in data.csv

**Step 3:** show a scatter plot of ‘Clump thickness vs No of week’ and plot their class labels given in column ‘Cancer stage’ (1, 2, 3, 4) for all 200 datapoints. See the reference output image below.

A graph of cancer stages

Description automatically generated

**Step 4:** Now, Plot the other added 2000 datapoints given in columns ‘Clump thickness\_new’ and ‘No of week\_new’ over the previous scatter plot without using their class label. See the reference output image below.

A diagram of cancer stages

Description automatically generated

**Step 5:** Now train the semi supervised model for the labeled 200 datapoints and make predictions for class label of new 2000 added datapoints and plot the scatters. Use KNN for base estimator. Use Model.fit() and model.predict() for training and prediction.

A diagram of a number of colored dots

Description automatically generated

**Step 6:** Now your added 2000 new data have their predicted class and true class are given in column ‘True cancer stage’. Compute and print accuracy score, plot classification report and the confusion matrix.

A graph of confusion matrix

Description automatically generated

DELIVERABLES

1. Submit code.
2. Scatter plots for all the output images above.

QUESTIONS

1. Provide three applications of Semi-supervised learning.
2. What are the three assumptions of Semi-supervised learning?
3. What is significance each of using the above three assumptions in Semi-supervised Learning?
4. How does the Co-training method differ from the Self-training method of Semi-supervised learning? Which one usually performs better for accuracy and such performance metrics?
5. How to evaluate the performance of a semi-supervised learning method while training?