1. **10 Points**

   Create a model to optimize prediction of the IRIS dataset using a perceptron.

   a. Load the IRIS dataset and split the data Split the data into two with 50% for training and 50% for testing.

   b. Display the first four rows of the dataset showing the feature names.

   c. Create a Perceptron class and instantiate a new Perceptron. Fit the data to the model for 10 training iterations. Compute the prediction.

   d. Use the first three features of the data and create a 3-D Scatter plot of the three types of flowers.

   e. Plot the prediction for epochs 1, 10, 100, and 200.

2. **10 Points**

Using the IRIS dataset again. Build a multilayer perceptron with a depth of 2, 20% input drop rate, and 5-% hidden layer drop rate. For the non-linear activation function, use the leaky rectify function for the first dense leayer, and softmax for the second dense layer.

Leaky Rectify Nonlinearity Function
The Leaky Rectifier has a non-zero gradient for negative inputs which often helps convergence:

$$\phi(v) = \begin{cases} v & \text{if } v \geq 0 \\ \alpha \cdot v & \text{otherwise.} \end{cases} \quad (1)$$

The slope for the negative input is α, a value between 0 and 1 which is a measure of "leakiness". A leakiness of 0 will converge to the standard rectifier, and a leakiness of 1 will lead to a linear activation function.

a. Plot the performance for the first 10 epochs.

b. Use a different activation function. What is the result?

c. Add more layers to the multilayer perceptron. Use 4 layers instead of 2. What is the result?

3. **10 Points**

   Load the Store Data file provided on Pilot and create a data frame of the list of items in the file.

   a. For the 7500 records, implement the Association Rule with a minimum support of 0.0045, minimum confidence of 0.2, and minimum lift of 3.

   b. Print the first 10 Rules along with each Support, Confidence and Lift values.

   c. How long did it take to run the implementation for the 7500 items.

   d. What are some other datasets for which this type of algorithm will be useful.

4. **10 Points**

   Open Orange 3 either from the command line or from the Anaconda Navigator. Using the IRIS dataset again, create a scatter plot of the data.

   a. What are the correlation values of all the combinations?

   b. What are the evaluation results for SMV, Random forest, and Logistic Regression.

   c. Show the confusion matrix for the three models.

   d. Repeat the process for zoo.tab dataset.

5. **10 Points**

   Load the breast cancer dataset from Scikit-Learn.

   a. What are the 13 features of this dataset?

   b. What are the labels of the data?

   c. Split the data into 70% training and 30% testing.

   d. Create a SVM classifier and train the model.

   e. Predict the output using the testing data.

   f. What is the accuracy, precision, and recall scores?

   g. Load this data into Orange and visualize the dataset and find informative projections.

6. **10 Points**

Using the IRIS dataset again, use the decision tree classifier to fit the data and plot the tree.

a. Display the decision surface of the tree for the first two features.

> **Decision Surface**
> *The Decision surface in a statistical classification problem, is the boundary of a hypersurface that partitions the underlying vector space of the data in each class. It shows where the algorithm separates the data of each each class.*

b. Is there a difference in the results from using SVM? What are the advantages and disadvantages when compared to SVM for this dataset?

c. Try displaying the decision surface using different features. Is there a difference? Is there better separation?

b. Use Orange to do the same.

7. **10 Points**

   In Orange, use the Breast Cancer dataset with a Random Forest model.

   a. Visualize the results using the Pythagorean Forest.

   b. What pre-processing did you do to reduce the data?

   c. What are the features that produce the best results? What can you use to make this simpler?

   d. How do you interpret the results?