



ASSIGNMENT 1

PERCEPTRON

CISC/CMPE 452/COGS400



General Instructions for Code and Submission (for all assignments)

1. You can use any programming language (preferred Python, java, C, C++, Java – **No GUI tool should be used**)
2. Make **one zip file named as Asgn1_studentID** which should include
 - a. *A PDF file (template given on the next page) containing the ANN design, accuracy in precision, recall, and confusion matrix*
 - b. *Program code with comments in the code to explain what each of your program files and functions are for*
3. Upload zip file to the OnQ site
4. Mark will be deducted based on late policy (-2 per day after due date until the end date after which no assignments will be accepted)

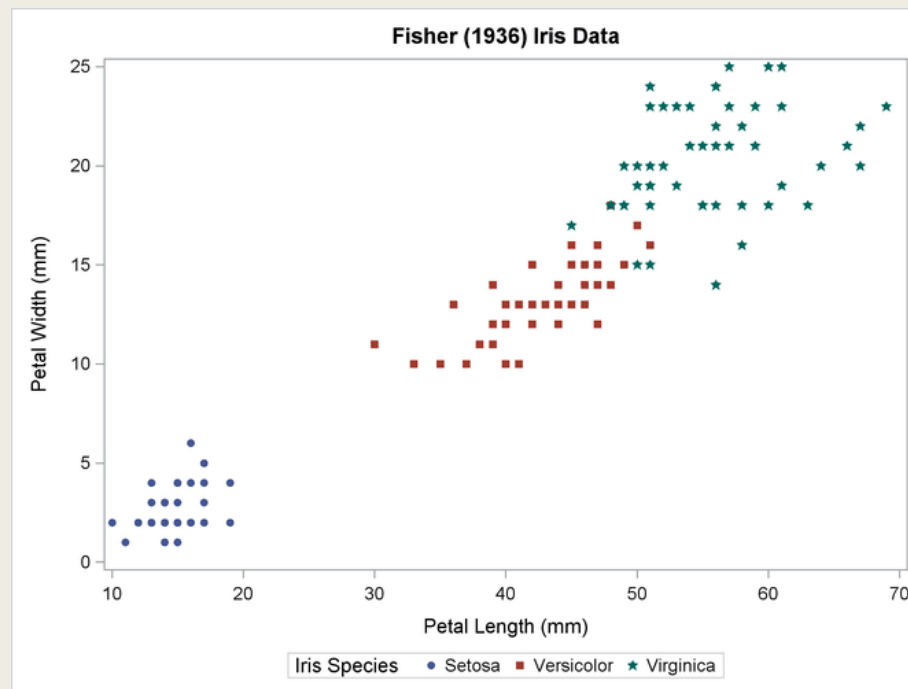
PDF file to be submitted (2 marks)

- For each part of the assignment list the following.
 1. Precision, Recall and overall accuracy in a confusion matrix with predicted class on y-axis and target class (actual class) on x-axis (see lecture slides for details). (1.5 marks)
 - a. *When data points are not linearly separable, they may be classified into more than one class. List these data points. Minimum error is indicated by fewest number of points wrongly classified.*
 2. Explanation of data preprocessing and the design decision of number of input, hidden and output nodes, and learning rate. (0.5 marks)

The Data

- Iris data is one of the most popular datasets in machine learning tasks. Iris is a flowering plant with 260-300 different species.
- Iris data contains the measurements of 3 different species of Iris: Setosa, Versicolour and Virginia.
- The plot on the next slide shows the relation between Petal width and length.
- The data for this assignment is provided in two files: iris_train.txt and iris_test.txt.
 - *Each line in the files represents a separate data point where each data point includes the following comma separated values:*
 - Sepal length (cm), Sepal width (cm), Petal length (cm), Petal width (cm), class label
 - Example: 5.1,3.5,1.4,0.2,Iris-setosa
 - *iris_train.txt contains 40 data points for each class. Use this to train your perceptron.*
 - *iris_test.txt contains 10 data points for each class. Use this to test that your perceptron works correctly.*

Plot of Iris data (Petal width and length)



Assignment 1 – Total 10 marks

- Design, program (no GUI tool should be used) and train a perceptron from scratch (do not use perceptrons implemented in other libraries) using the data in iris_train.txt to predict
 - **Part A** - Predict the class labels of all the data points using simple feedback learning ($\Delta w = cx$). (5 marks – 3 for fully functional and correct executable and code with comments + 2 for report)
 - **Part B** – Apply Pocket algorithm and report the differences in performance with Part A. (5 marks – 3 for for fully functional and correct executable and code with comments + 2 for report)
 - Use iris_test.txt to test the accuracy of your ANN.