**Course No.: ELEN-857**

**Course Title: Advanced Pattern Recognition Method**

**Department: Electrical and Computer Engineering**

**Project 1: Data creation and display**

**Submitted To: Prepared By:**

**Dr. Robert Y.Li, Professor Name: Mrinmoy Sarkar**

**Department of Electrical Engineering Banner Id: 950-363-260 Telephone: (336) 285-3716; E-mail: eeli@ncat.edu E-mail: msarkar@aggies.ncat.edu**

**Contents:**

1. **Abstract**
2. **Technical Description**
3. **Results**
4. **Summary**
5. **Appendix**
6. **Abstract:**

The main purpose of the project is to load data from a file and manage the data effectively to use in future project to implement different pattern recognition algorithm. And to visualize the data by plotting in a 2-d graph. For this project Fisher’s Iris data is used which contains a set of measurements related to 3 species of the Iris plant. The three species are Iris Setosa, Iris Versicolor, and Iris Virginica.

1. **Technical Description:**

The dataset contains 50 plants from each of the 3 species. There are 4 features in the dataset named sepal length, sepal width, petal length and petal width. MATLAB programming language is used for loading the data from the file and plotting the data.

1. **Results:**

(1) 2-d plot of the Fisher’s Iris data with petal width versus sepal length

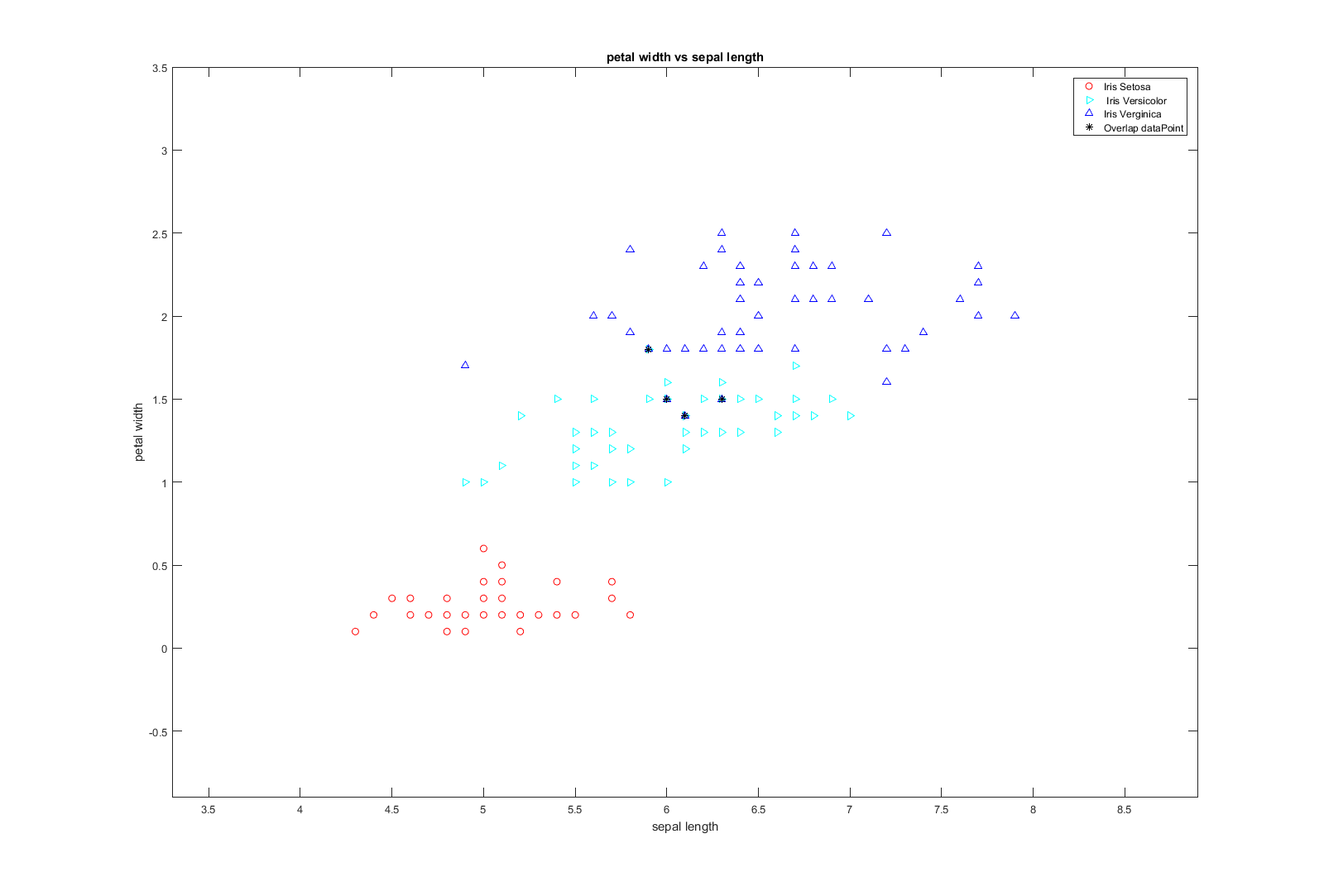


Figure 1: 2-d plot of the Fisher’s Iris data with petal width versus sepal length

# Total overlapping data points for petal width vs sepal length is: 4

(2) 2-d plot of the Fisher’s Iris data with petal length versus sepal width

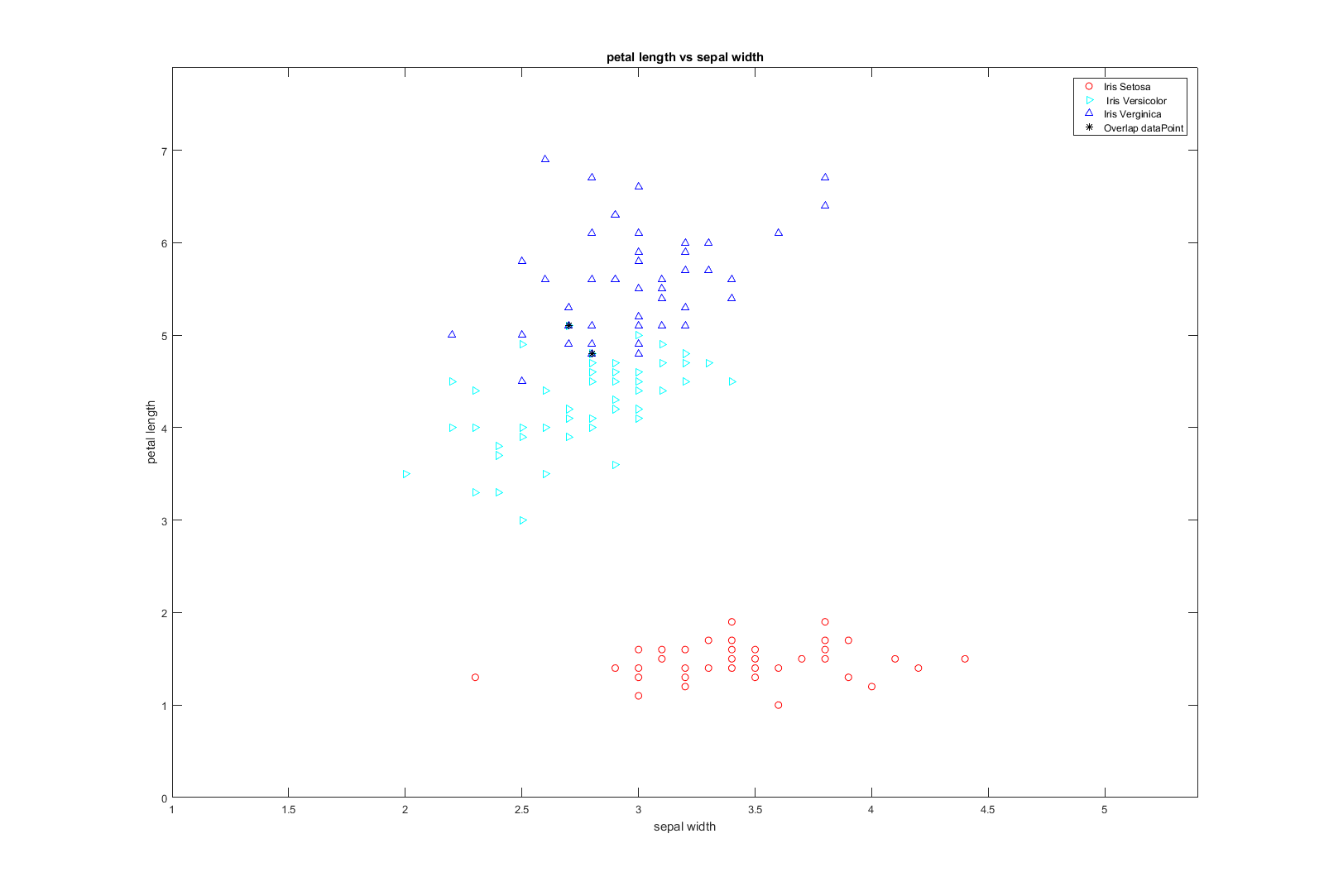


Figure 2: 2-d plot of the Fisher’s Iris data with petal length versus sepal width

# Total overlapping data points for petal length vs sepal width is: 2

1. **Summary:**

**(1)**

* For petal width vs sepal length, Iris Setosa is linearly separable from Irsi Vesicolor and Iris Verginica
* There are 4 overlapping points between Irsi Vesicolor and Iris Verginica
* Because of the overlapping points, Irsi Vesicolor and Iris Verginica are not linearly separable

**(2)**

* For petal length vs sepal width, Iris Setosa is linearly separable from Irsi Vesicolor and Iris Verginica
* There are 2 overlapping points between Irsi Vesicolor and Iris Verginica
* Because of the overlapping points, Irsi Vesicolor and Iris Verginica are not linearly separable
* There are some points in Irsi Vesicolor and Iris Verginica that are enclosed in one another region

1. **Appendix**

**MATLAB Code:**

%% file name dataCreationAndDisplay.m

% author: Mrinmoy Sarkar

% email: msarkar@aggies.ncat.edu

% date: 9/2/2017

clear;

close all;

% load data to a veriable

data = importdata('iris.txt');

% no. of class is 3 named Iris-setosa, Iris-versicolor and Iris-verginica

% there are 4 attributes named sepal-length, sepal-width, petal-length,

% petal-width

% there are 50 plants for each species

irisSetosa = zeros(50,4);

irisVersicolor = zeros(50,4);

irisVerginica = zeros(50,4);

n = size(data,1);

indxSeto = 1;

indxVers = 1;

indxVerg = 1;

for i=2:n

x = strsplit(cell2mat(data(i)));

if strcmp(x(5), 'Iris-setosa')

for j=1:4

irisSetosa(indxSeto,j) = str2double(cell2mat(x(j)));

end

indxSeto = indxSeto + 1;

elseif strcmp(x(5), 'Iris-versicolor')

for j=1:4

irisVersicolor(indxVers,j) = str2double(cell2mat(x(j)));

end

indxVers = indxVers + 1;

elseif strcmp(x(5), 'Iris-virginica')

for j=1:4

irisVerginica(indxVerg,j) = str2double(cell2mat(x(j)));

end

indxVerg = indxVerg + 1;

end

end

%% plot Fisher's Iris data with petal width versus sepal length

figure;

plot(irisSetosa(:,1), irisSetosa(:,4), 'or'); hold on;

plot(irisVersicolor(:,1), irisVersicolor(:,4), '>c'); hold on;

plot(irisVerginica(:,1), irisVerginica(:,4), '^b'); hold on;

axis([min([min(irisSetosa(:,1)),min(irisVersicolor(:,1)),min(irisVerginica(:,1))])-1 ...

max([max(irisSetosa(:,1)),max(irisVersicolor(:,1)),max(irisVerginica(:,1))])+1 ...

min([min(irisSetosa(:,4)),min(irisVersicolor(:,4)),min(irisVerginica(:,4))])-1 ...

max([max(irisSetosa(:,4)), max(irisVersicolor(:,4)),max(irisVerginica(:,4))])+1]);

title('petal width vs sepal length');

xlabel('sepal length');

ylabel('petal width');

legend('Iris Setosa', ' Iris Versicolor', 'Iris Verginica','AutoUpdate','off');

%% plot overlap symbol (\*)

x1 = table(irisSetosa(:,1),irisSetosa(:,4));

x2 = table(irisVersicolor(:,1),irisVersicolor(:,4));

x3 = table(irisVerginica(:,1),irisVerginica(:,4));

temp1 = union(intersect(x1,x2),intersect(x1,x3));

overlapingPoints = union(temp1, intersect(x2,x3));

disp(['Total overlapping data points for petal width vs sepal length is : ' num2str(size(overlapingPoints,1))])

plot(overlapingPoints.Var1, overlapingPoints.Var2, '\*k');

legend('Iris Setosa', ' Iris Versicolor', 'Iris Verginica', 'Overlap dataPoint')

%% plot Fisher's Iris data with petal length versus sepal width

figure;

plot(irisSetosa(:,2), irisSetosa(:,3), 'or'); hold on;

plot(irisVersicolor(:,2), irisVersicolor(:,3), '>c'); hold on;

plot(irisVerginica(:,2), irisVerginica(:,3), '^b'); hold on;

axis([min([min(irisSetosa(:,2)),min(irisVersicolor(:,2)),min(irisVerginica(:,2))])-1 ...

max([max(irisSetosa(:,2)),max(irisVersicolor(:,2)),max(irisVerginica(:,2))])+1 ...

min([min(irisSetosa(:,3)),min(irisVersicolor(:,3)),min(irisVerginica(:,3))])-1 ...

max([max(irisSetosa(:,3)), max(irisVersicolor(:,3)),max(irisVerginica(:,3))])+1]);

title('petal length vs sepal width');

xlabel('sepal width');

ylabel('petal length');

legend('Iris Setosa', ' Iris Versicolor', 'Iris Verginica','AutoUpdate','off');

%% plot overlap symbol (\*)

x1 = table(irisSetosa(:,2),irisSetosa(:,3));

x2 = table(irisVersicolor(:,2),irisVersicolor(:,3));

x3 = table(irisVerginica(:,2),irisVerginica(:,3));

temp1 = union(intersect(x1,x2),intersect(x1,x3));

overlapingPoints = union(temp1, intersect(x2,x3));

disp(['Total overlapping data points for petal length vs sepal width is : ' num2str(size(overlapingPoints,1))])

plot(overlapingPoints.Var1, overlapingPoints.Var2, '\*k');

legend('Iris Setosa', ' Iris Versicolor', 'Iris Verginica', 'Overlap dataPoint')