

Name: Mrinmoy Sarkar
Banner Id: 950-363-260

Project 1: Data creation and display

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

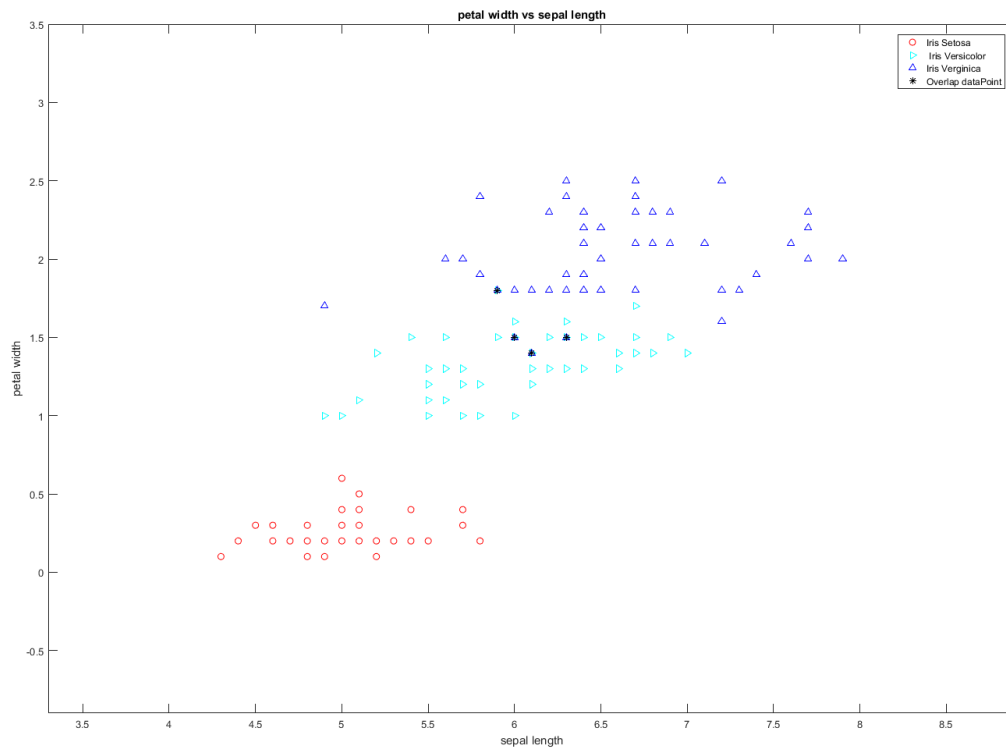


Figure 1: 2-d plot of the Fisher's Iris data with petal width versus sepal length

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

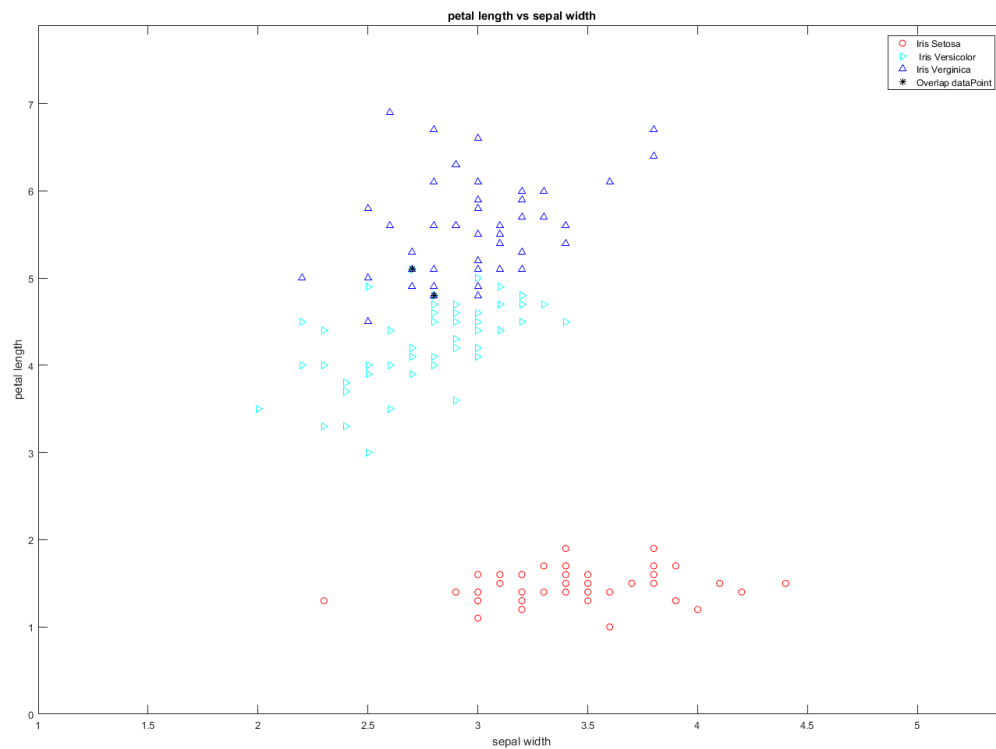


Figure 2: 2-d plot of the Fisher's Iris data with petal length versus sepal width

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

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 variable
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 ...
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));
overlappingPoints = union(temp1, intersect(x2,x3));

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

plot(overlappingPoints.Var1, overlappingPoints.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));
overlappingPoints = union(temp1, intersect(x2,x3));

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

plot(overlappingPoints.Var1, overlappingPoints.Var2, '*k');
legend('Iris Setosa', ' Iris Versicolor', 'Iris Verginica', 'Overlap
dataPoint')

```

Summary of the results:

(1)

- For petal width vs sepal length, Iris Setosa is linearly separable from Iris Versicolor and Iris Virginica
- There are 4 overlapping points between Iris Versicolor and Iris Virginica
- Because of the overlapping points, Iris Versicolor and Iris Virginica are not linearly separable

(2)

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