

# مبانی داده‌کاوی درخت تصمیم (مثال)

محمد جواد فدائی اسلام

## مجموعه داده گل‌های IRIS



Iris setosa



Iris versicolor



Iris virginica

- از هر نوع گل ۵۰ نمونه انتخاب شد و طول و عرض گلبرگ (petal) و کاسبرگ (sepal) آنها اندازه‌گیری شد.
- پس ۱۵۰ داده (نمونه) داریم که از هر داده ۴ ویژگی وجود دارد.

## پایگاه داده

Meas (measurements) (features)				Species (Class)
SL (Sepal Length)cm	SW (Sepal Width)cm	PL (Petal Length)cm	PW (Petal Width)cm	Type
۵,۱	۳,۵	۱,۴	۰,۲	setosa
۴,۹	۳	۱,۴	۰,۲	setosa
۴,۷	۳,۲	۱,۳	۰,۲	setosa
۴,۶	۳,۱	۱,۵	۰,۲	setosa
۷	۳,۲	۴,۷	۱,۴	versicolor
...				

## دستور متلب FITCTREE

>>fitctree

Fit a classification decision tree.

`tree = fitctree(X,Y)`

returns a classification tree based on the input variables (also known as predictors, features, or attributes)  $X$  and output (response or labels)  $Y$ . The returned tree is a binary tree, where each branching node is split based on the values of a column of  $X$ .

## مثال

```
load fisheriris
t = fitctree(meas,species,'PredictorNames',{'SL' 'SW' 'PL' 'PW'});
view(t)
view(t,'mode','graph');
```

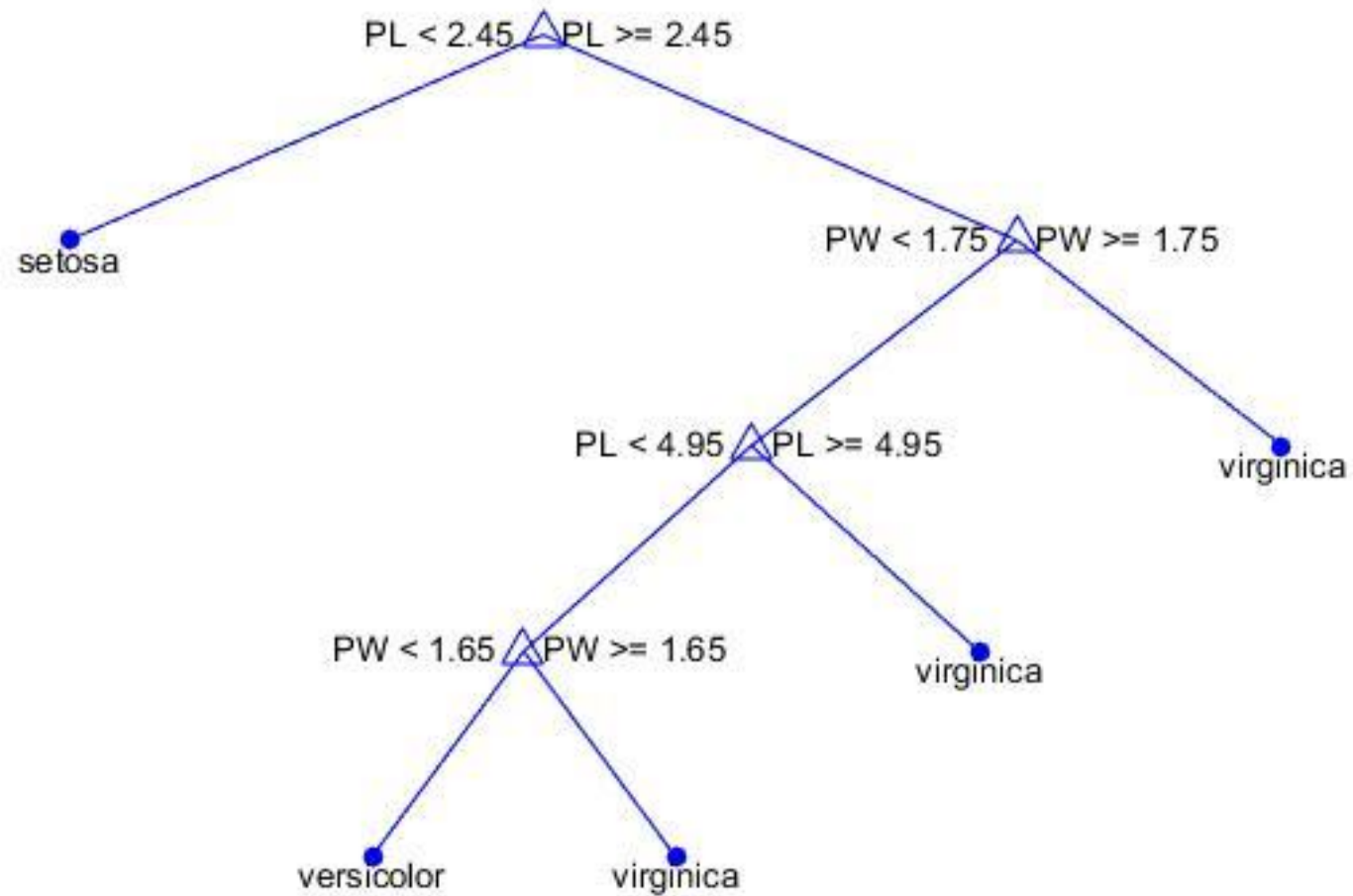
## >>VIEW(T)

Decision tree for classification

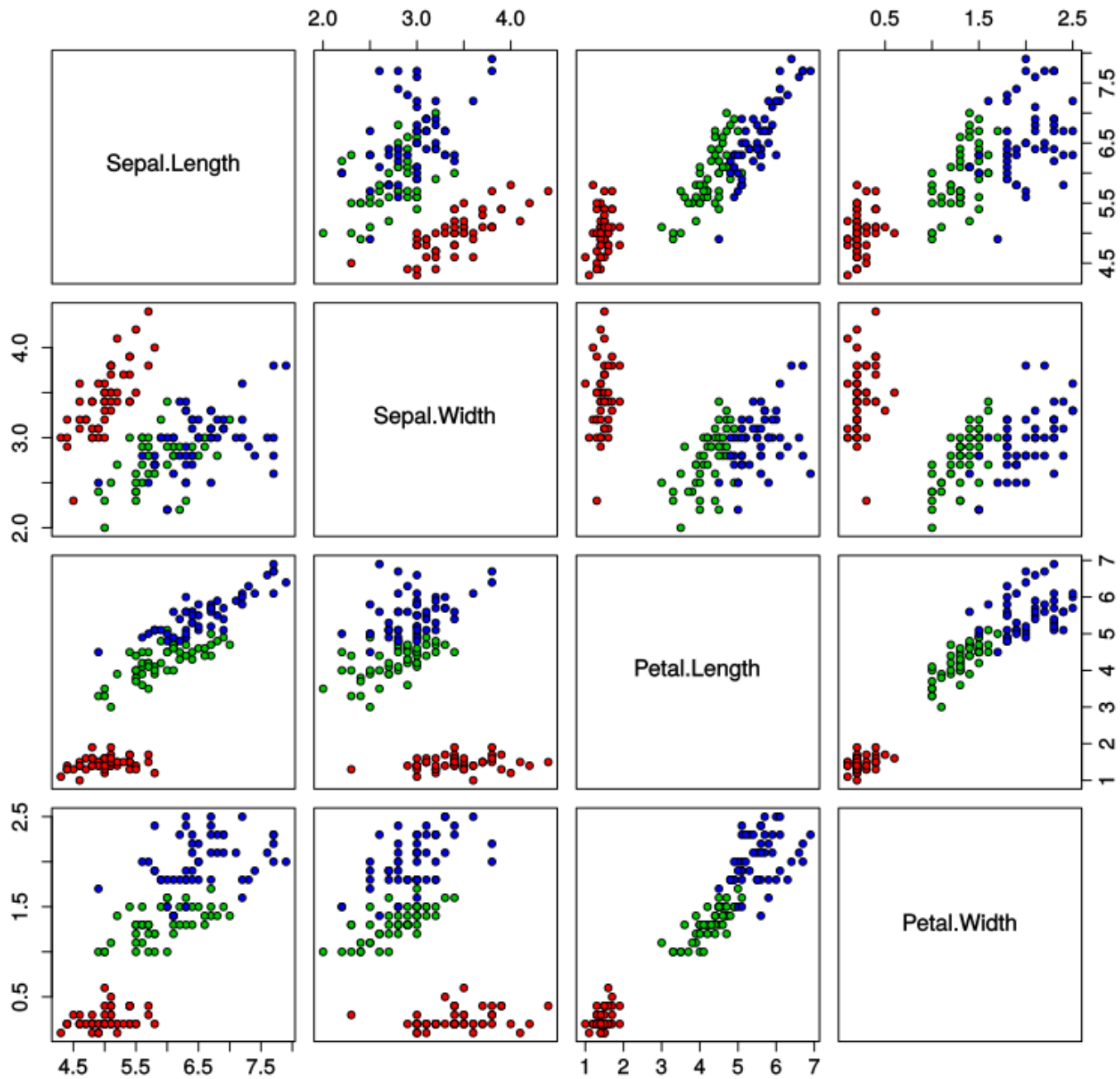
- 1 if  $PL < 2.45$  then node 2 elseif  $PL \geq 2.45$  then node 3 else setosa
- 2 class = setosa
- 3 if  $PW < 1.75$  then node 4 elseif  $PW \geq 1.75$  then node 5 else versicolor
- 4 if  $PL < 4.95$  then node 6 elseif  $PL \geq 4.95$  then node 7 else versicolor
- 5 class = virginica
- 6 if  $PW < 1.65$  then node 8 elseif  $PW \geq 1.65$  then node 9 else versicolor
- 7 class = virginica
- 8 class = versicolor
- 9 class = virginica

```
>>VIEW(T,'MODE','GRAPH');
```

---



# Iris Data (red=setosa,green=versicolor,blue=virginica)



# SCATTER PLOT



# SCATTER PLOT

```
figure; hold on;  
plot(meas(1:50,3),    meas(1:50,4),    '+r');  
plot(meas(51:100,3), meas(51:100,4), '+g');  
plot(meas(101:150,3), meas(101:150,4), '+b');  
xlabel('PL')  
ylabel('PW')  
title('Scatter plot of IRIS')
```

