

A practical guide to machine learning for health actuaries

Dave Liner

13 JUNE 2019



BE AN ACTUARY.

BROUGHT
TO YOU BY:



WHAT
IS AN
ACTUARY?

WHY
ACTUARIAL
SCIENCE?

HOW
DO I GET
STARTED?

COLLEGE
STUDY

ACTUARIAL
EXAMS

FINDING
A JOB

DIVERSITY
PROGRAMS

NEWS &
RESOURCES

WHAT IS AN ACTUARY?

▶ WHAT DO WE DO?

▶ WHERE DO WE WORK?

▼ A TOP RANKED JOB

- ARTICLES

▪ FAST FACTS ABOUT ACTUARIES

Top 10 Reasons to Consider Being an Actuary: #3

You want to be able to choose among outstanding job offers.

Related Links

- [Tips for Working Overseas](#)
- [Career Outlook for Actuaries](#)
- [Ask an Actuary](#)

Ranking the actuarial career.

- [25 High-Paying Jobs for People Who Don't Like Stress - Actuary](#)
- [Actuary makes Time's list of highest paying jobs in 2016.](#)
- [CareerCast Ranks Actuary as One of the Twelve Best Jobs for Women in 2014](#)
- [Top 10 STEM Careers - Actuarial Profession Profile](#)
- [Actuaries In Action: Why It's Rated The Number One Profession](#)
- [The Best Jobs of 2013 - Actuary #1](#)
- [Four degrees with 0% unemployment: Actuary job in hand, before the degree](#)
- [Careercast.com reported in the Wall Street Journal, actuaries have the second best job in 2012.](#)
- [Best Careers for 2012](#)
- [Reassuring News for Actuarial Science](#)
- [CNBC Highlights Actuary as the 3rd Best Job in America for 2011](#)
- [CareerCast Ranks Actuary #3 on its 2011 List of Best Jobs](#)
- [Actuary Named Best Job for 2010 as Reported in the Wall Street Journal Online](#)
- [Most Lucrative College Degrees](#)
- [Actuary Ranks as One of the Best Jobs During Recession](#)
- [Actuary #8 on List of Best Jobs for College Graduates](#)
- [Actuary is #2 in Ranking of Best Occupations in the U.S.](#)
- [Ten Jobs that Pay \\$80,000 a Year \(Actuary is #1\)](#)

← Share

CareerCast top job rankings by year

Rank	2009	2010	2011	2012	2013	2014	2015	2016	2017
1									
2									
3									
4									
5									
6									
7									
8									
9									
10									
11									

Actuary

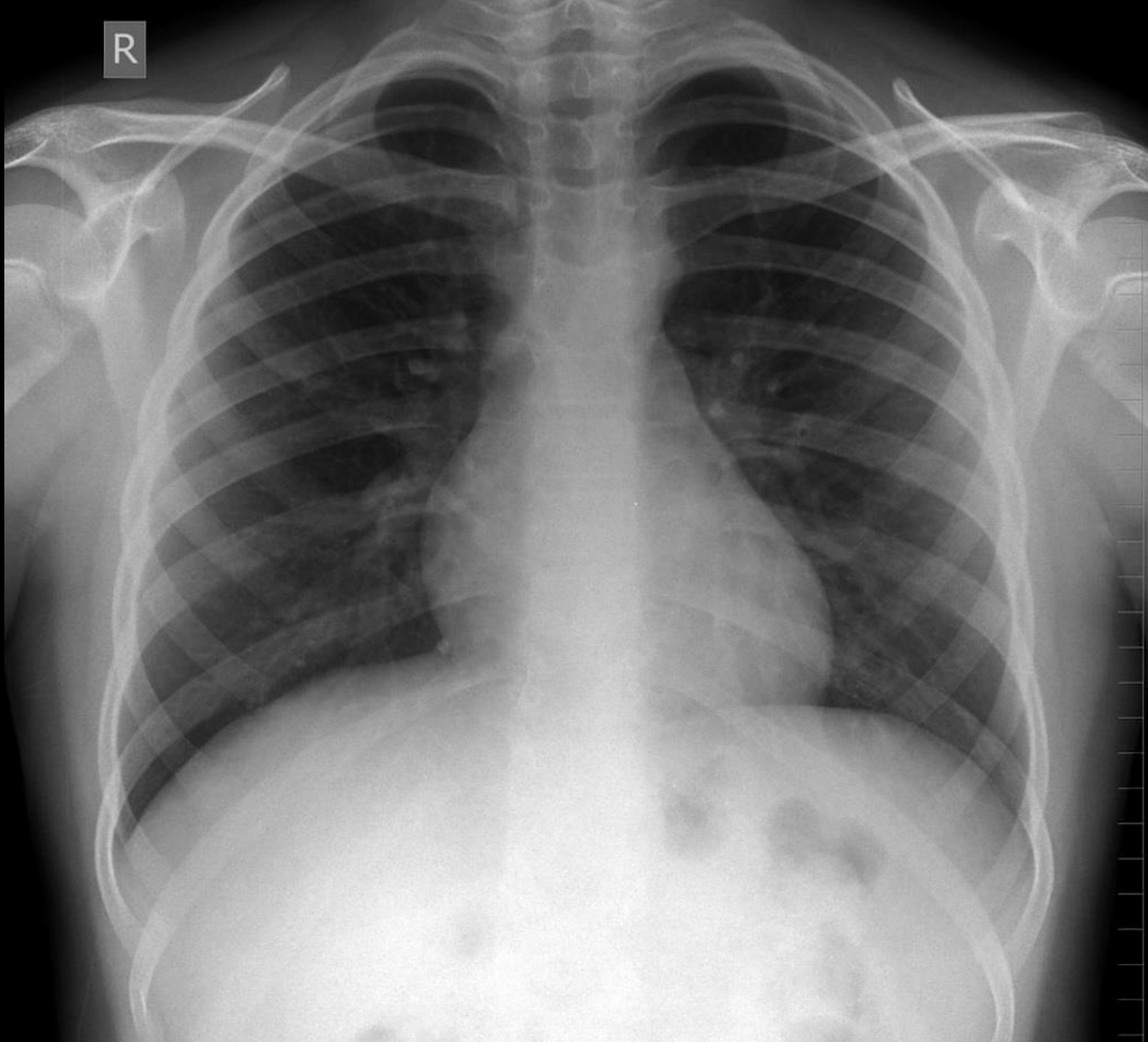
Data Scientist

Statistician

CareerCast top job rankings by year

Rank	2009	2010	2011	2012	2013	2014	2015	2016	2017
1		Actuary			Actuary		Actuary	Data Scientist	Statistician
2	Actuary			Actuary				Statistician	
3	Statistician		Actuary			Statistician			
4			Statistician			Actuary	Statistician		
5									Data Scientist
6							Data Scientist		
7									
8		Statistician							
9									
10								Actuary	
11									Actuary

R





What I learned about machine learning

1. Machine learning drives disruption and innovation in many sectors globally
2. Many actuaries already do most of the machine learning process
3. Most popular machine learning methods are based on concepts that actuaries already understand
4. There is a staggering amount of free resources to develop machine learning skills

Actuaries already do most of the machine learning process

1. Get data
2. Prepare data
3. Build model
4. Use model to gain insight
5. Tell others about model results

We will review five machine learning methods today

k-Nearest neighbors

k-Means clustering

Decision trees

Logistic regression

Neural networks

Illustrative dataset

Setosa

Versicolor

Virginica

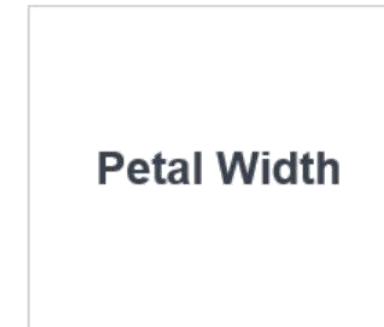
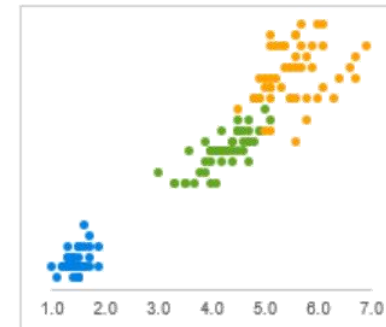
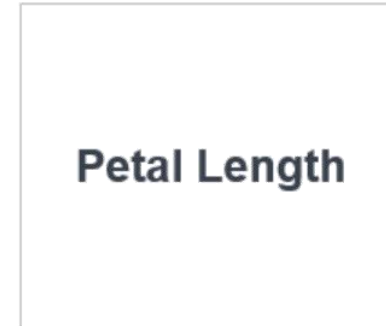
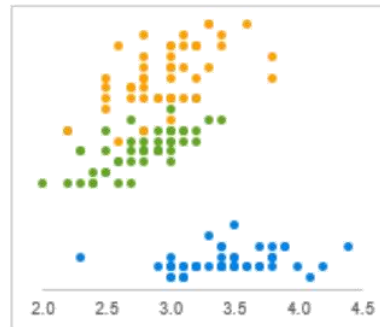
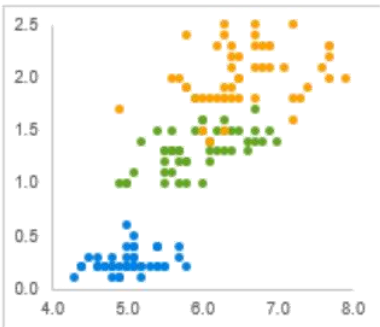
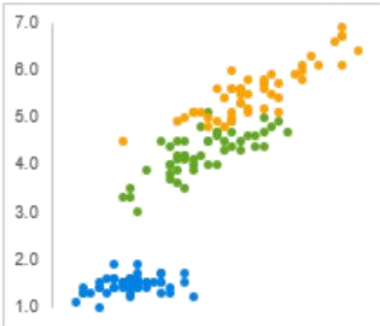
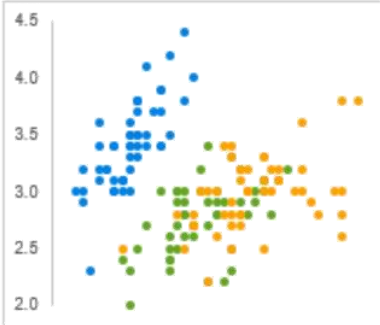
Sepal Length

Sepal Width

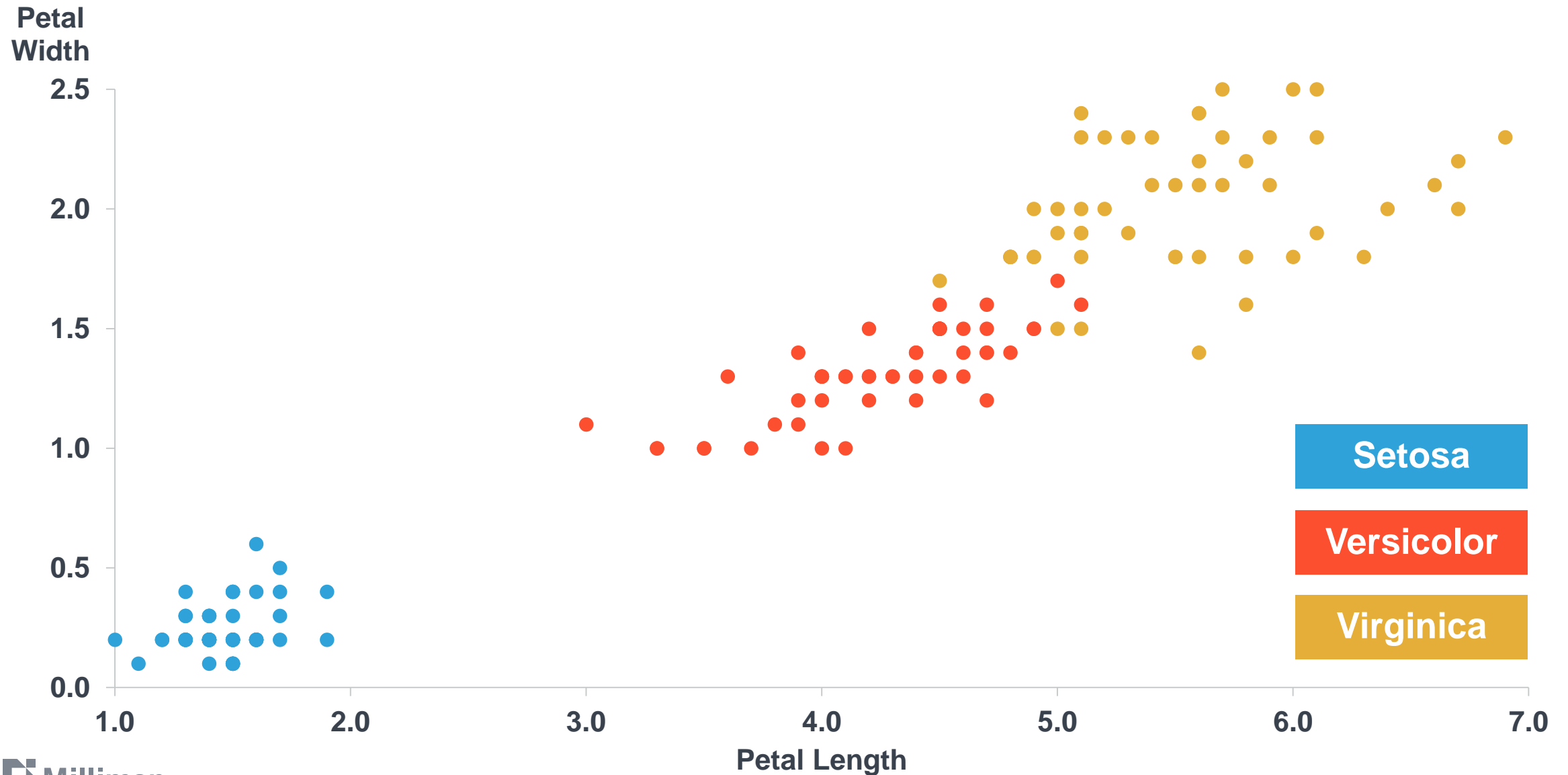
Petal Length

Petal Width

X1 = Petal length
X2 = Petal width
Y = Species

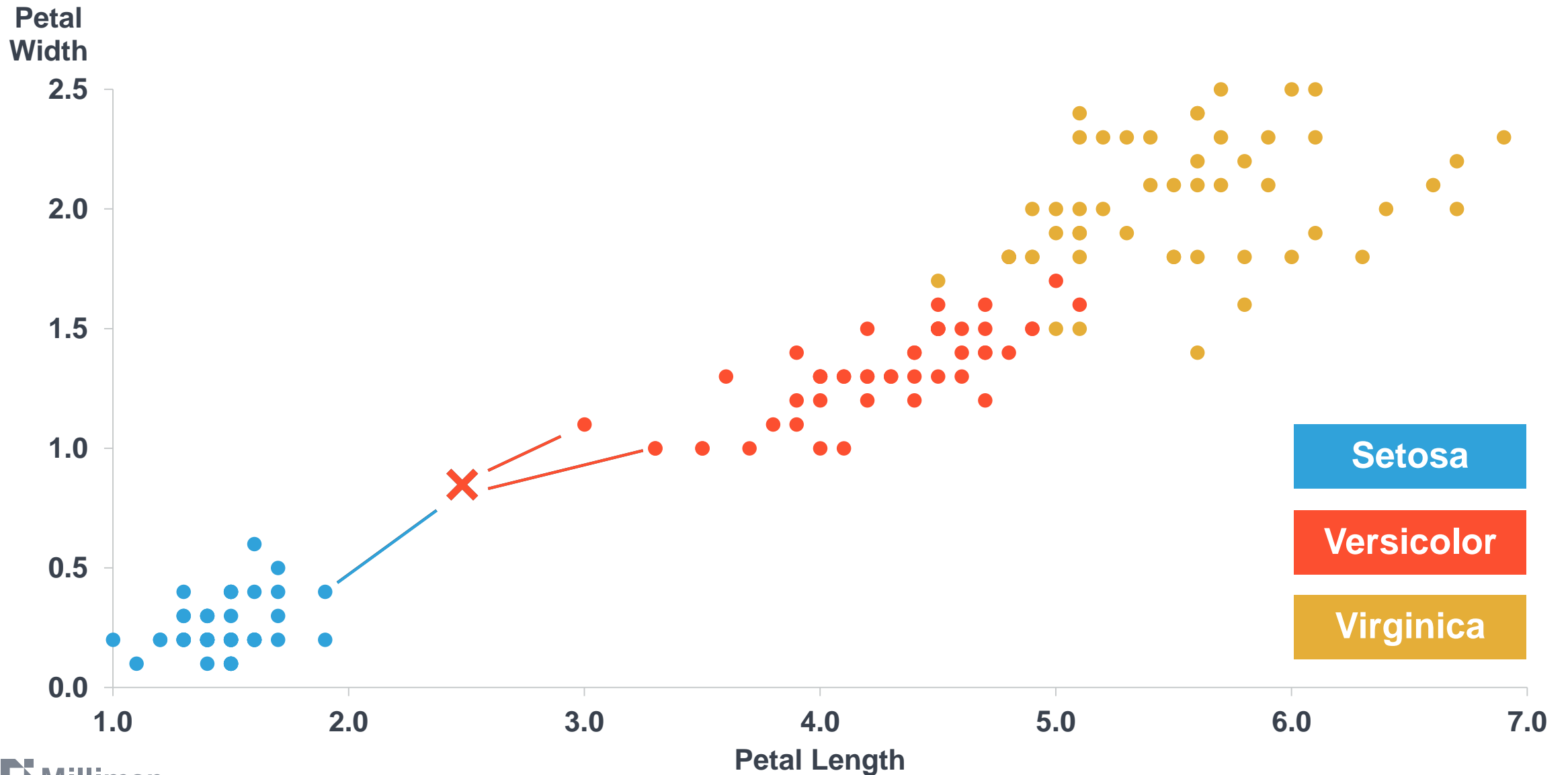


It is easier to illustrate methods using 2-dimensional data

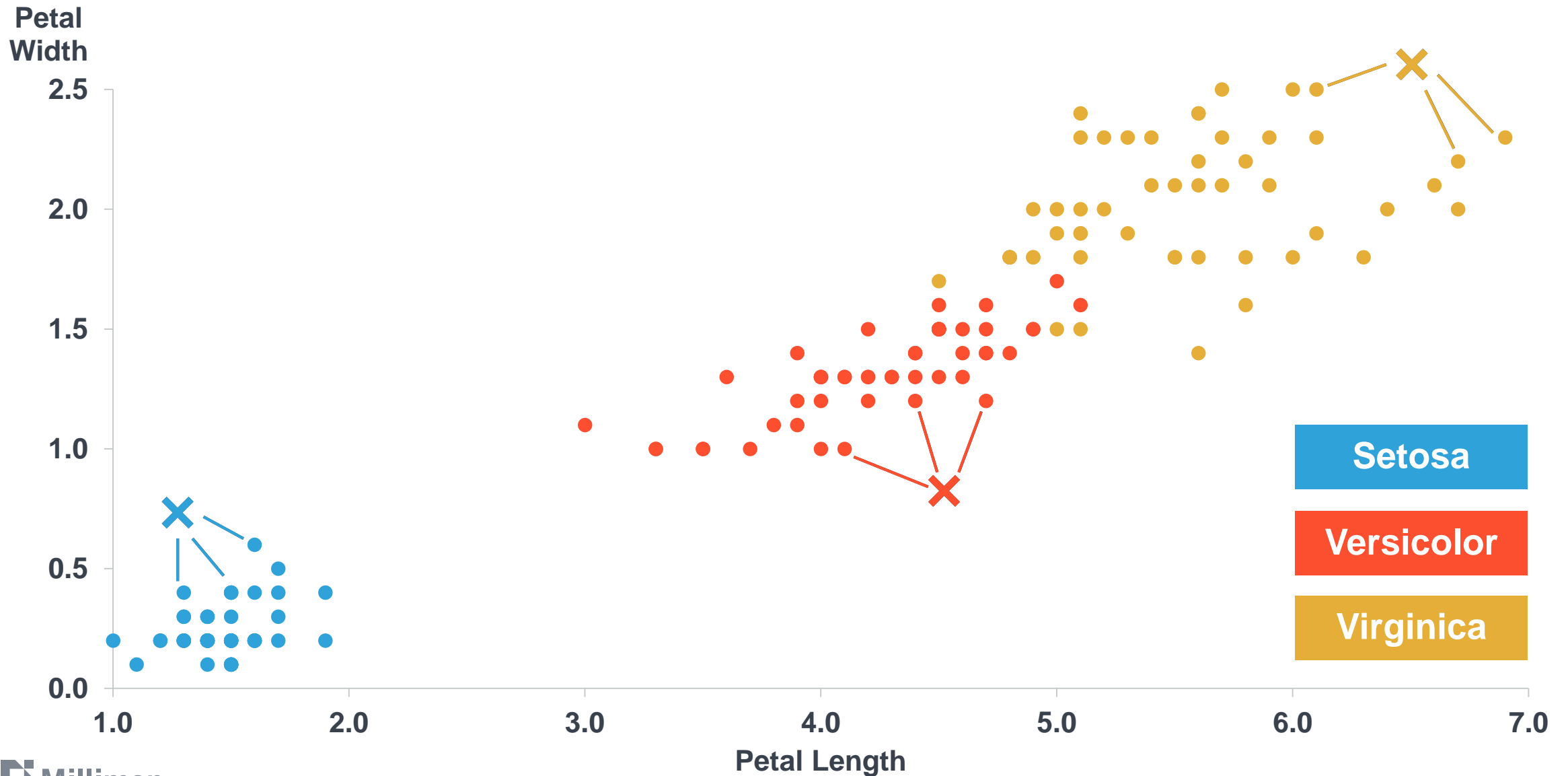


k-Nearest neighbors

What is the species of a new sample based on Petal width and length using kNN?



What is the species of a new sample based on petal width and length using kNN?

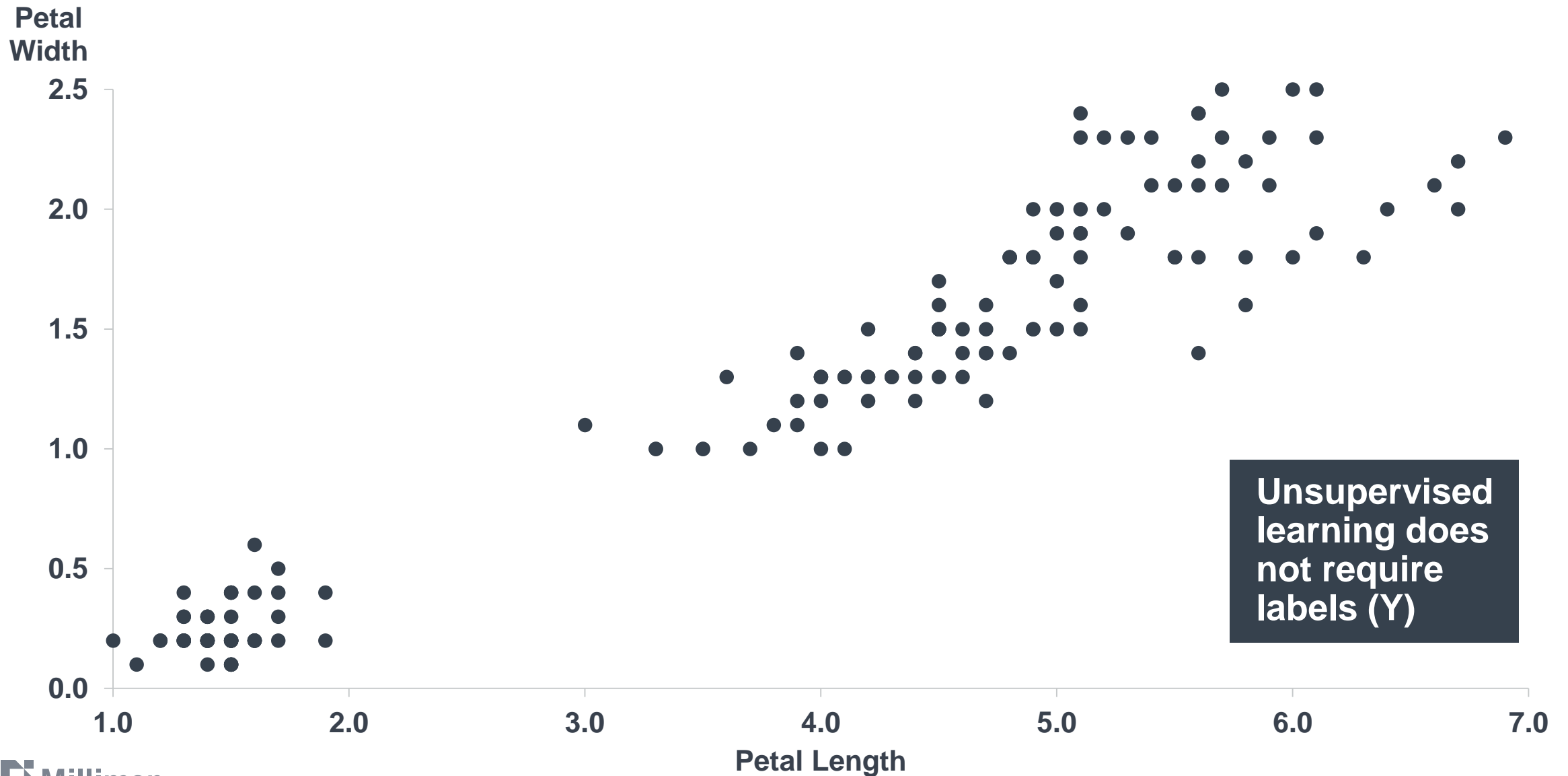


Four lines of Python code gets same result as Excel kNN file

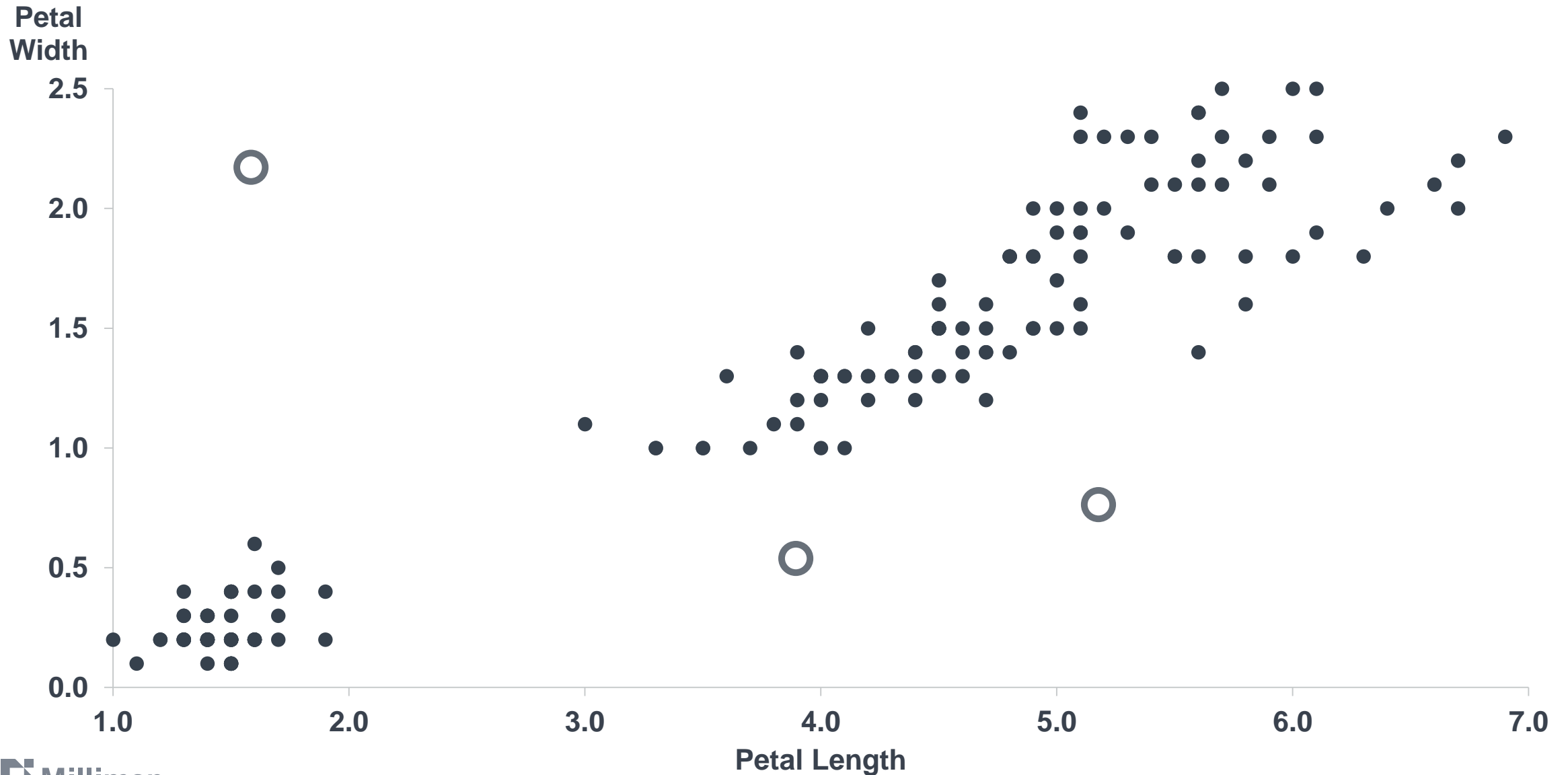
```
import sklearn.datasets as ds, sklearn.model_selection as ms, sklearn.preprocessing as pp, sklearn.neighbors, sklearn.metrics
X_train,X_test,y_train,y_test = ms.train_test_split(pp.scale(ds.load_iris().data),ds.load_iris().target,test_size=30,random_state=0)
y_pred_kNN = sklearn.neighbors.KNeighborsClassifier(n_neighbors=3).fit(X_train,y_train).predict(X_test)
print(sklearn.metrics.confusion_matrix(y_test, y_pred_kNN))
```

k-Means clustering

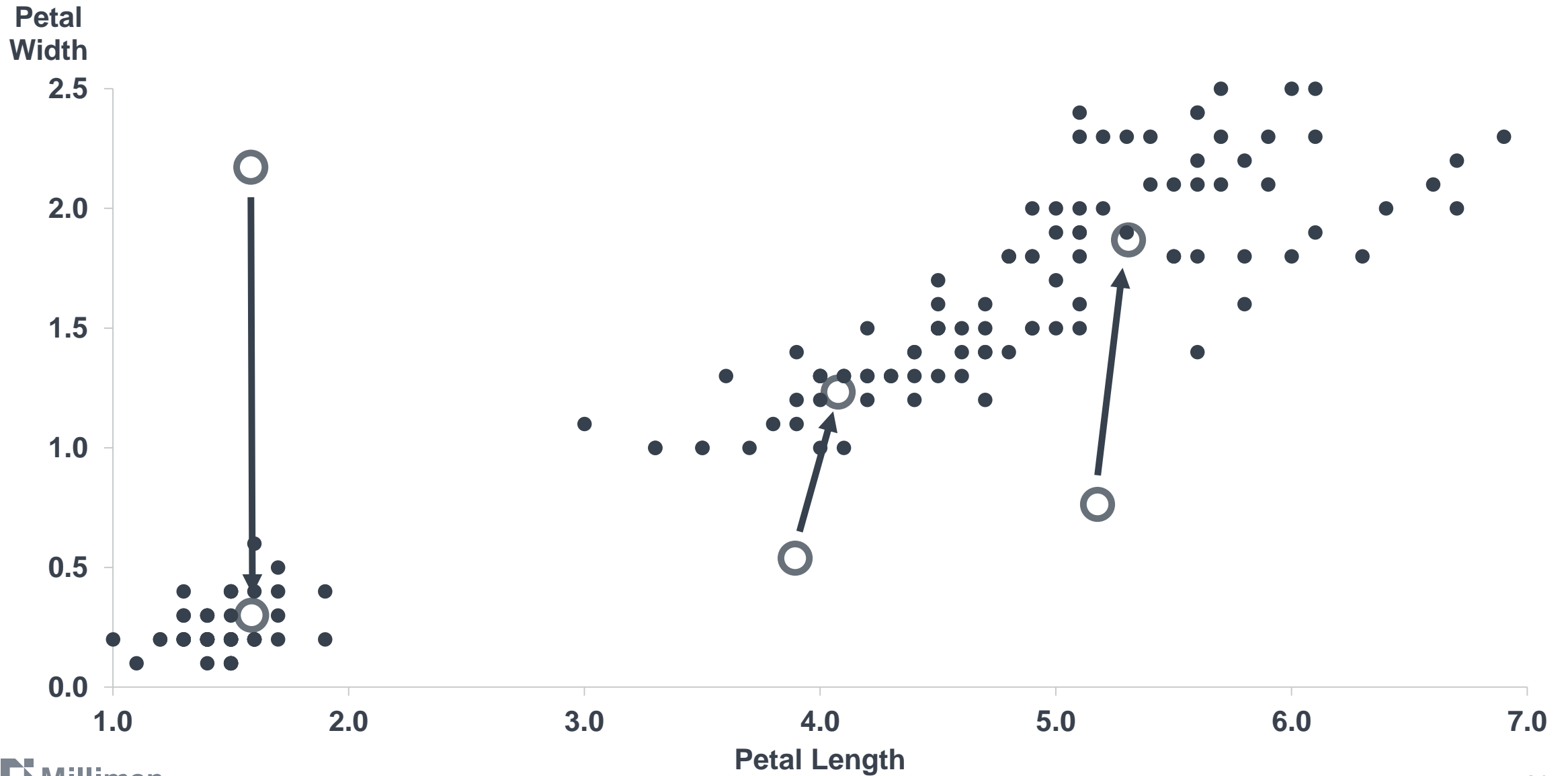
What is the species of a new sample based on Petal width and length using clustering?



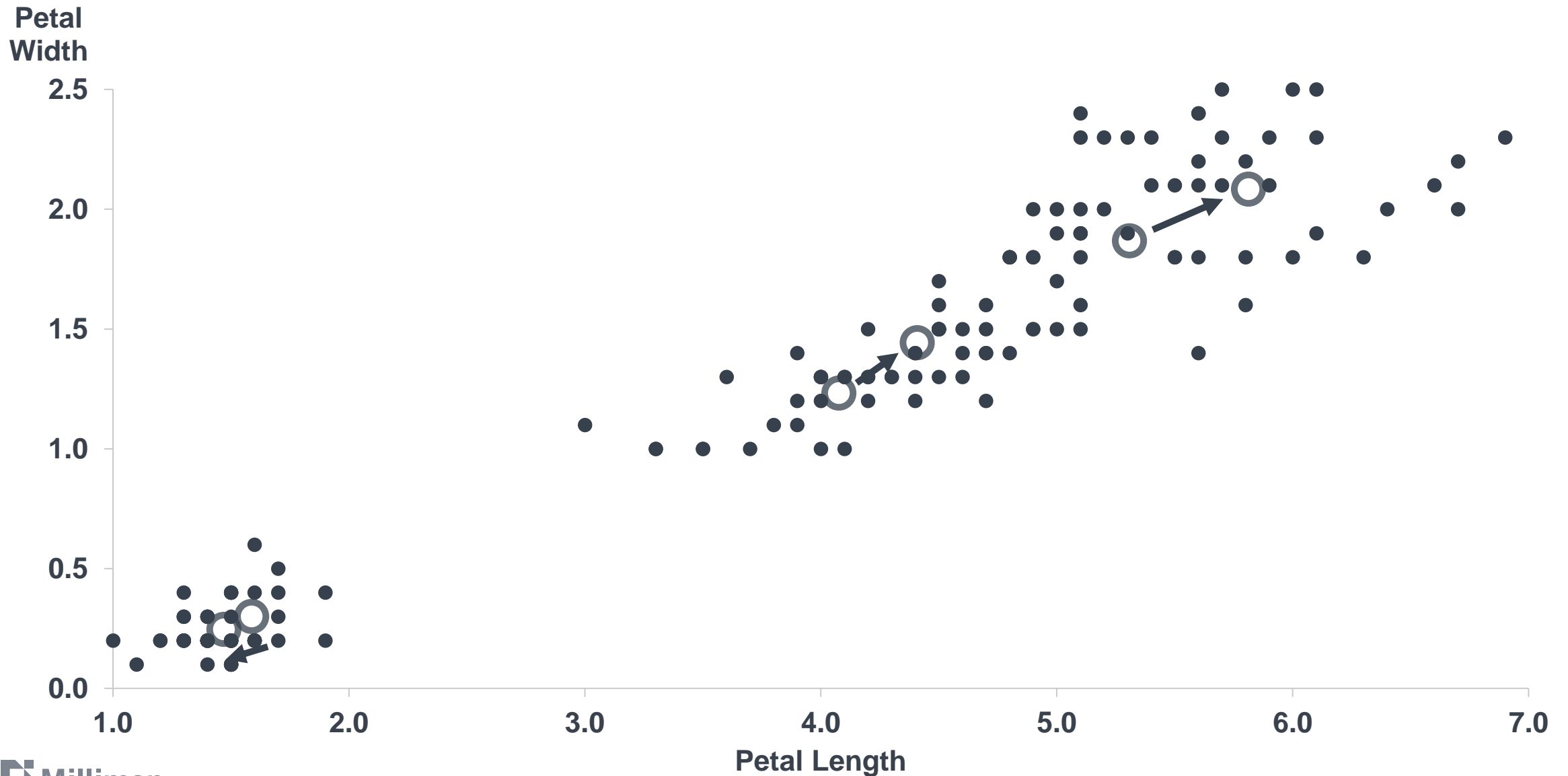
Step 1: select random centroids (k=3 in this case)



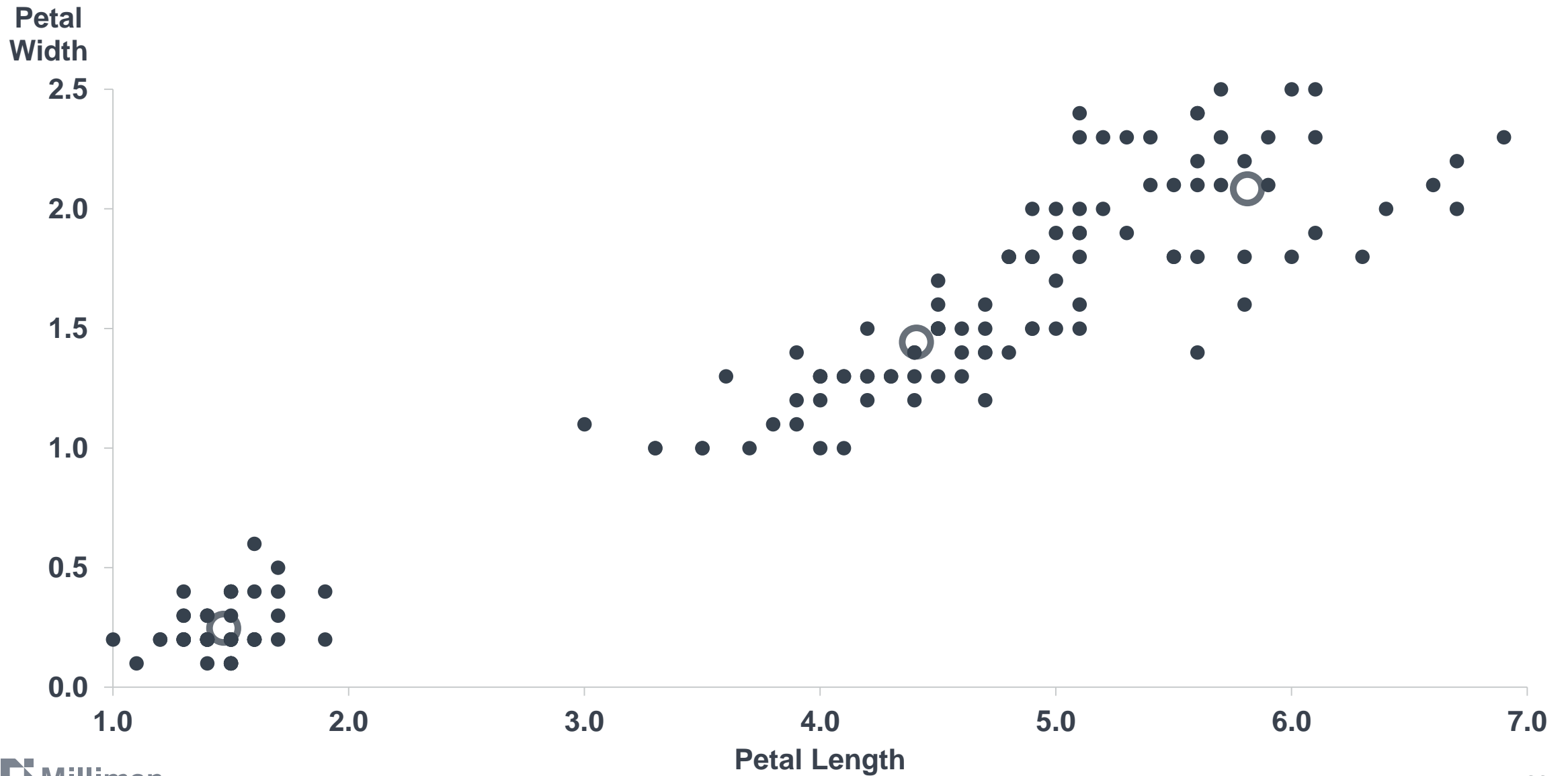
Step 2: move centroids based on data



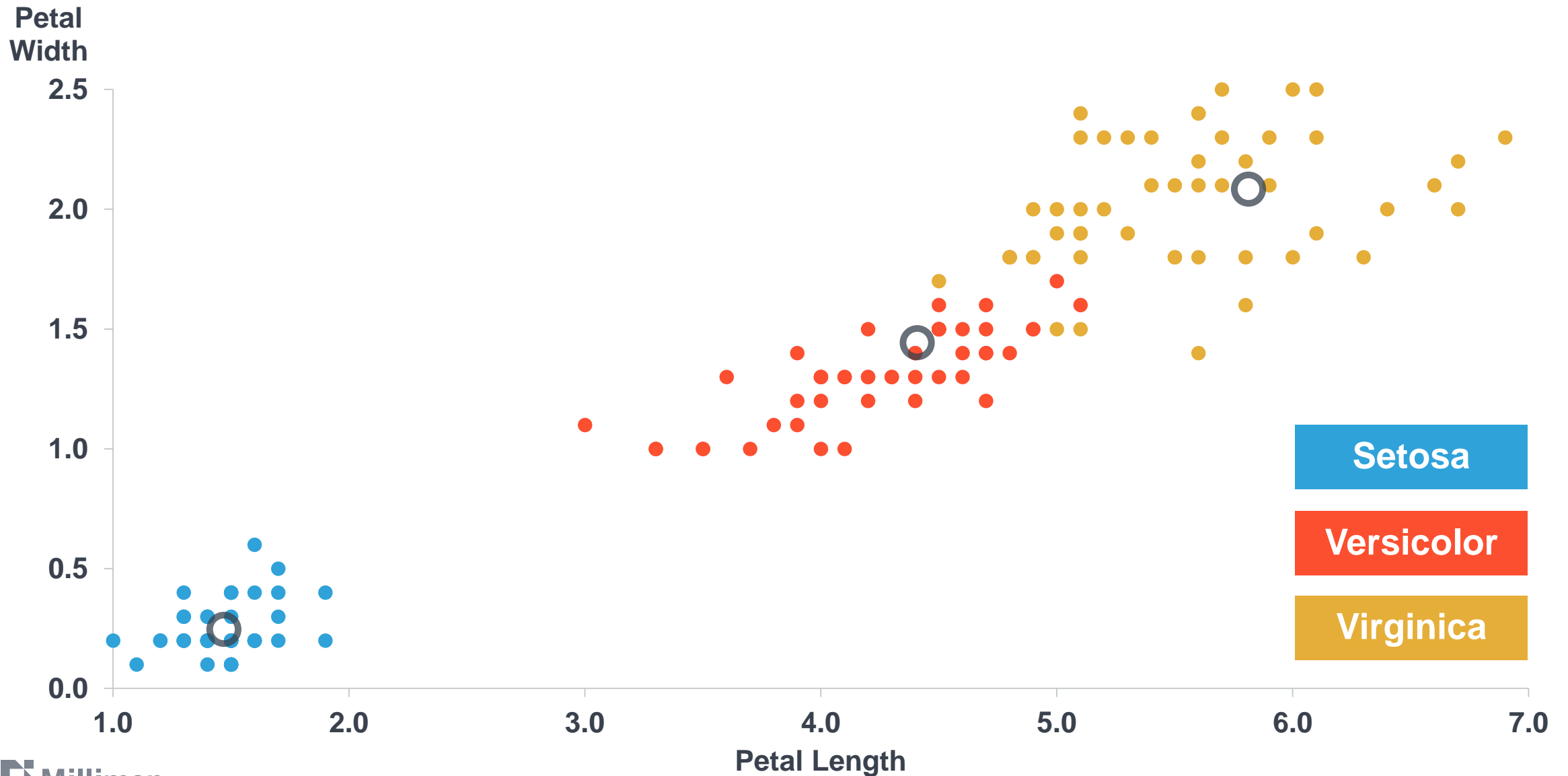
Step 3: keep moving centroids until no points are reassigned



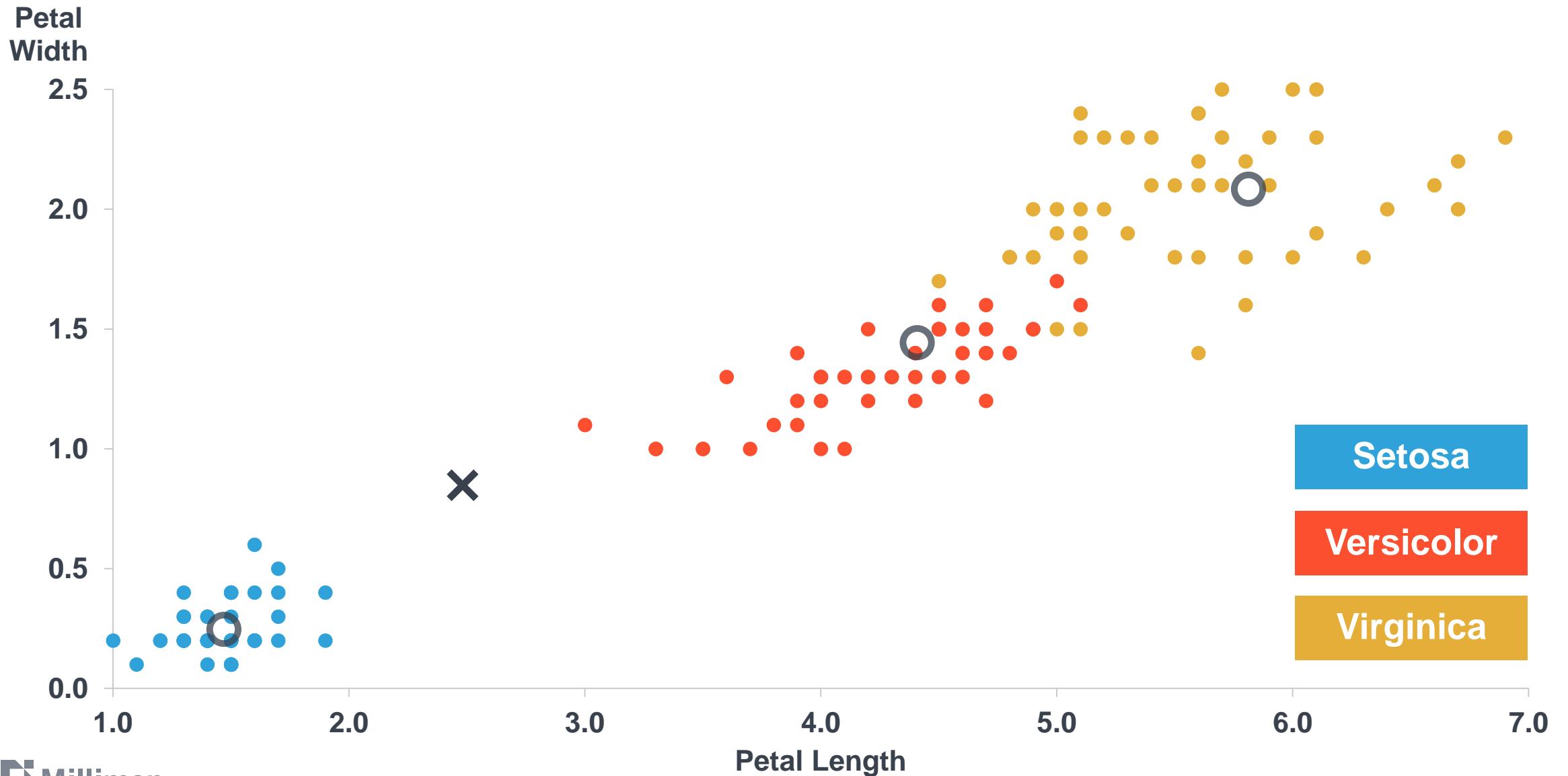
Step 4: assign each point to a centroid cluster



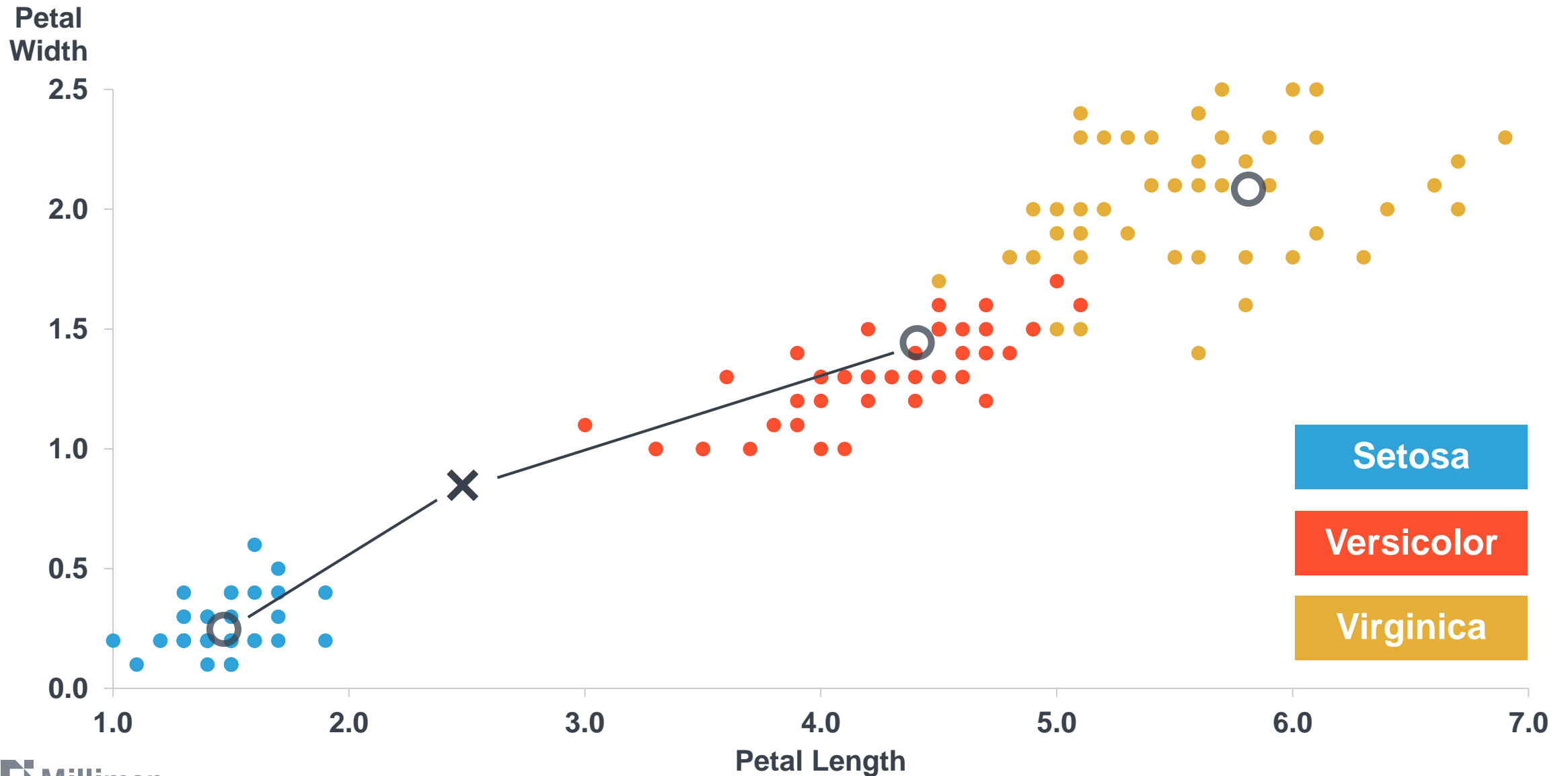
Step 4: assign each point to a centroid cluster



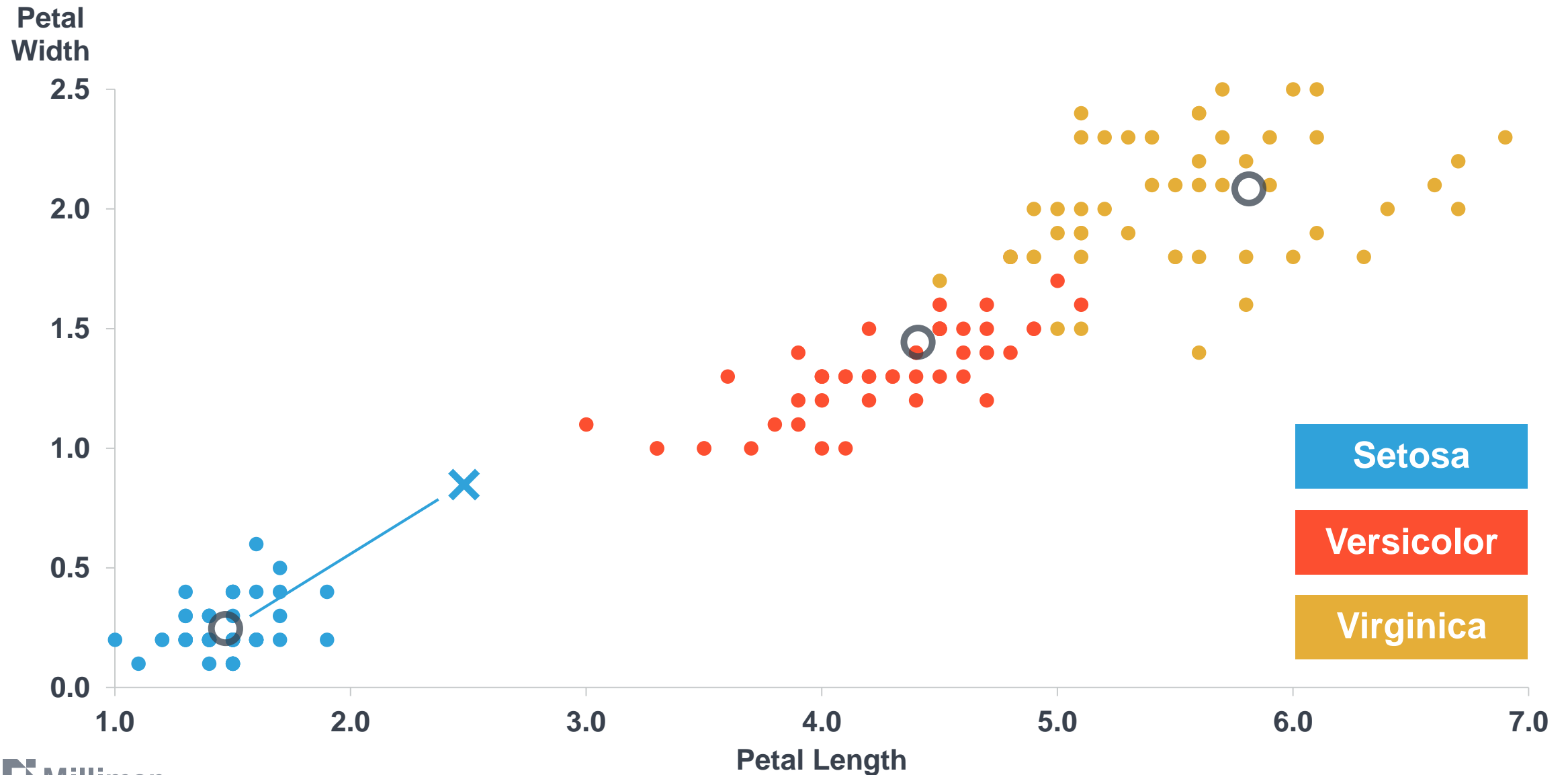
Step 5: use clusters to assign new data points



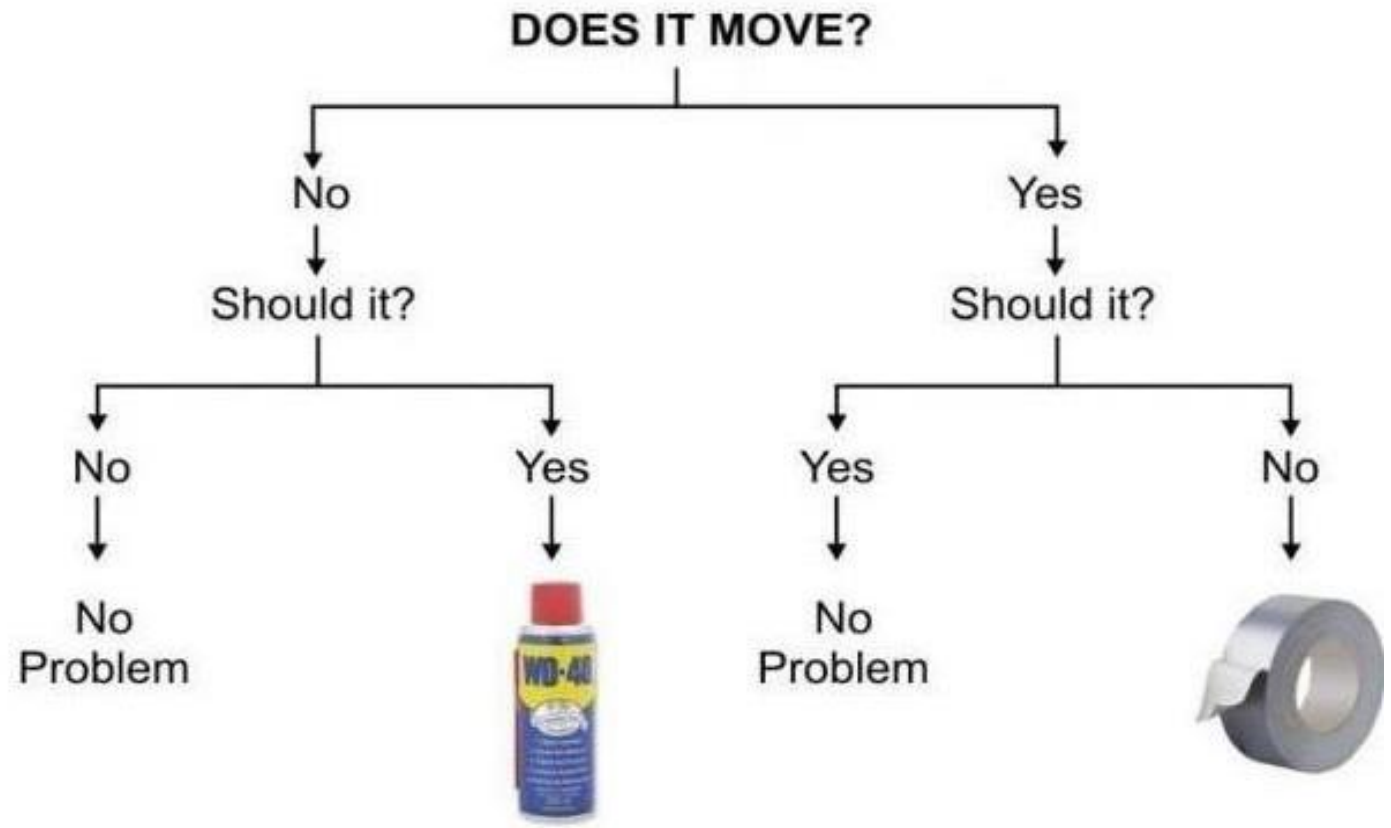
Step 5: use clusters to assign new data points



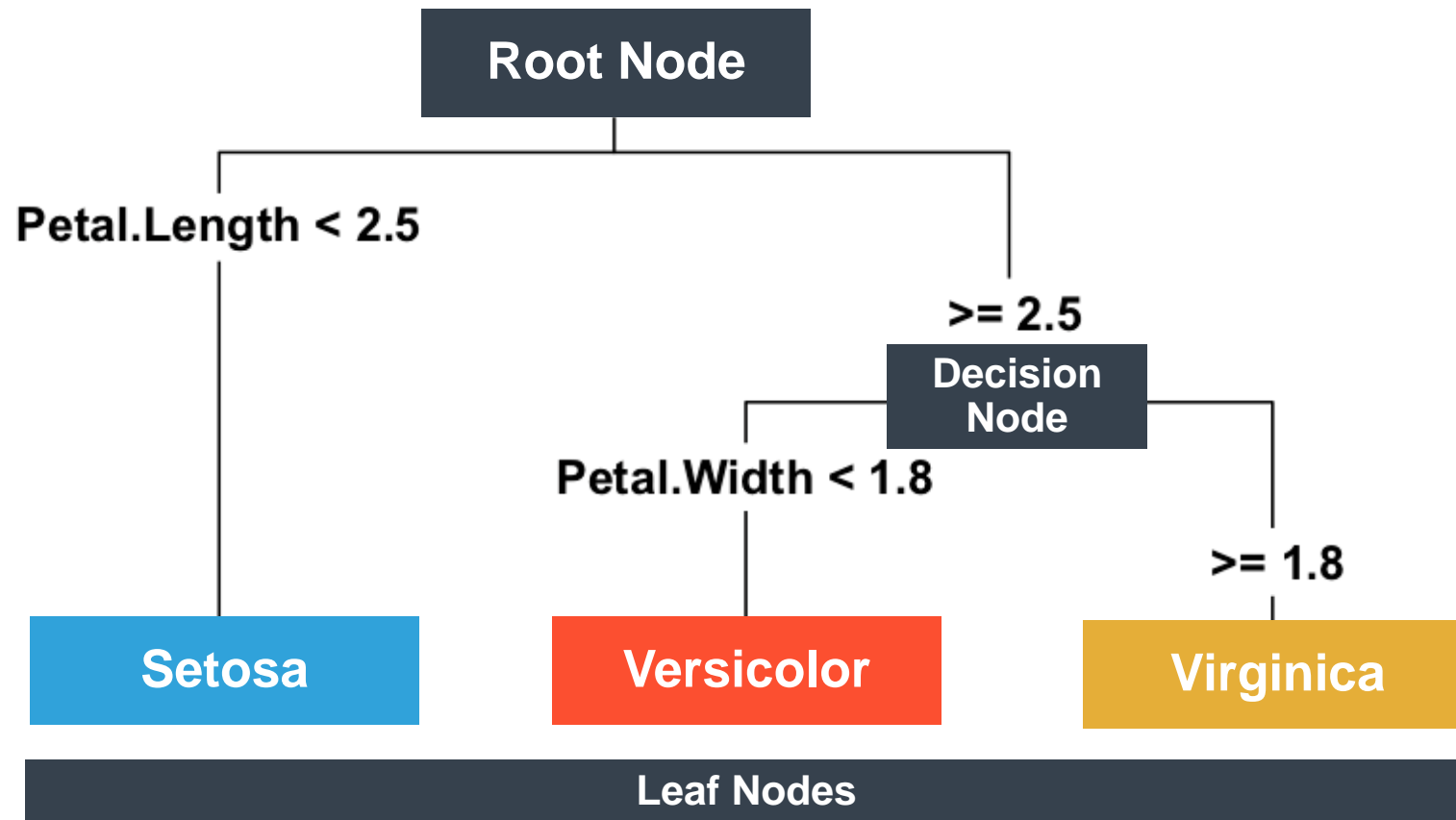
Step 5: use clusters to assign new data points



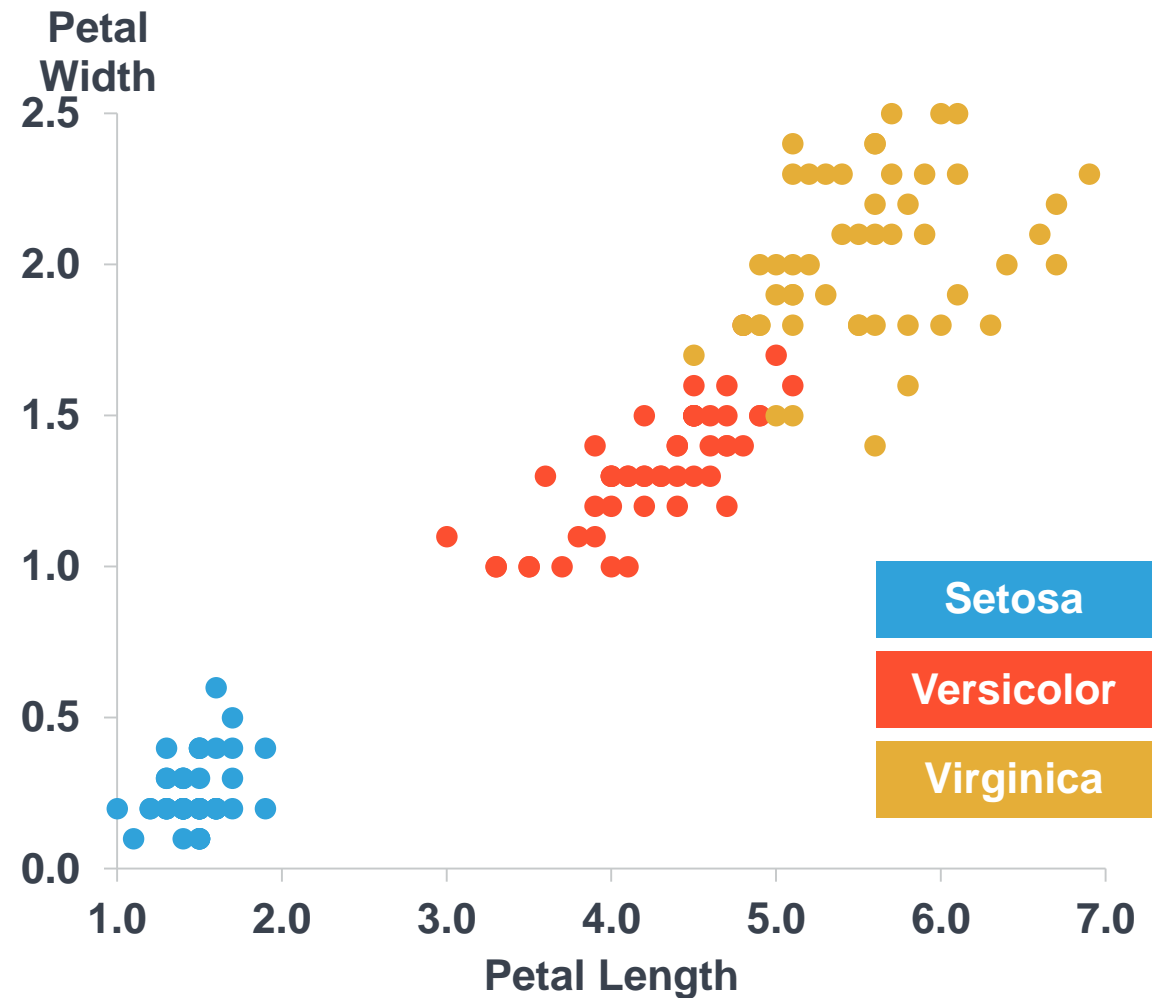
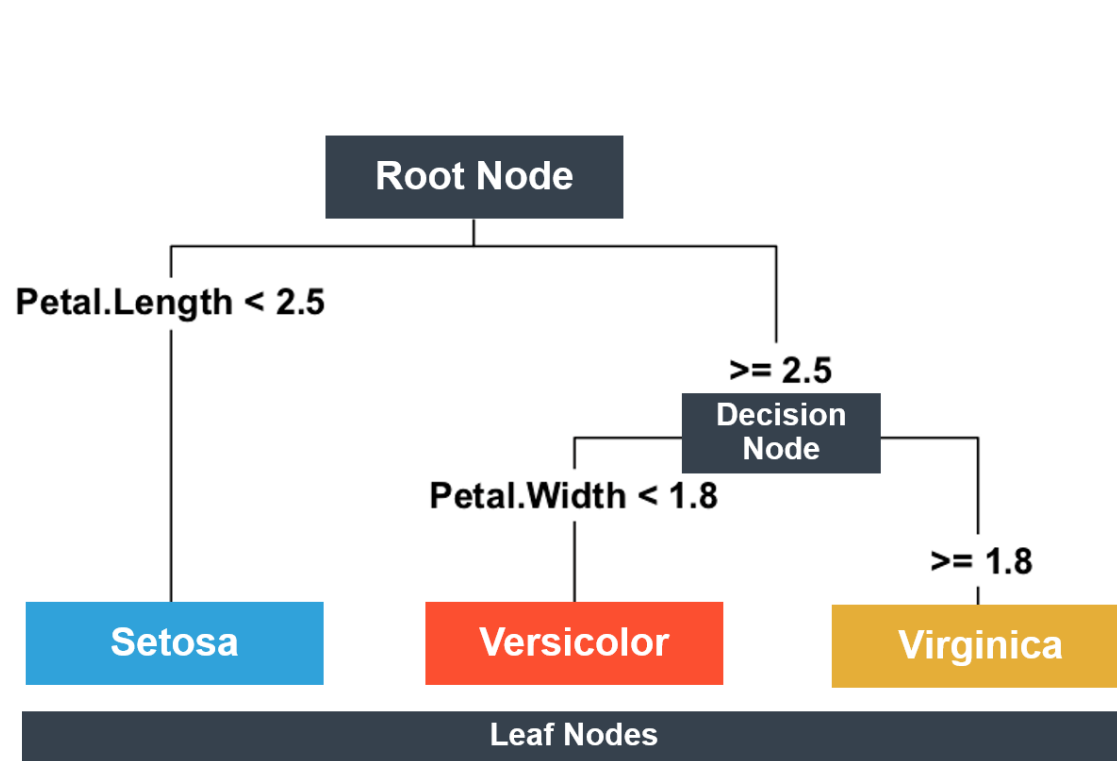
Decision trees



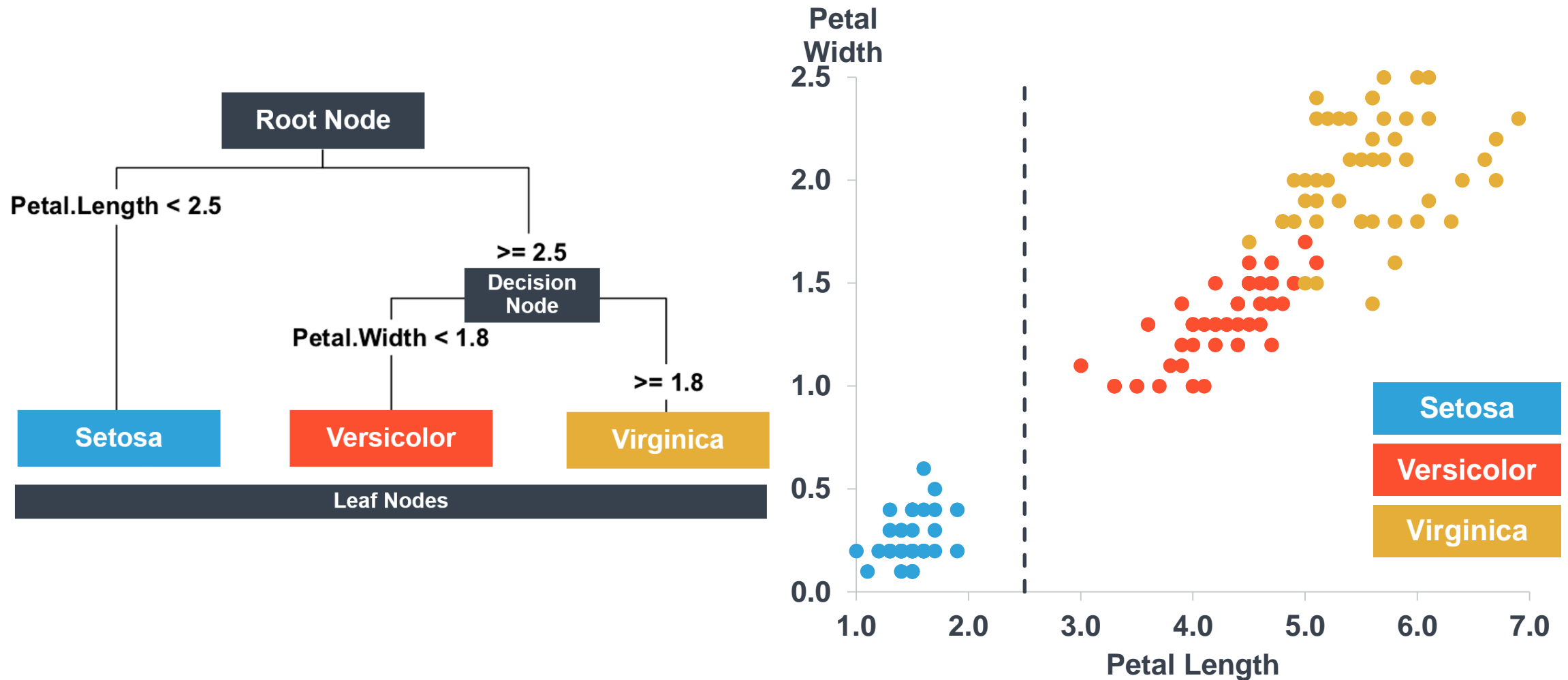
Decision trees provide another method for classification



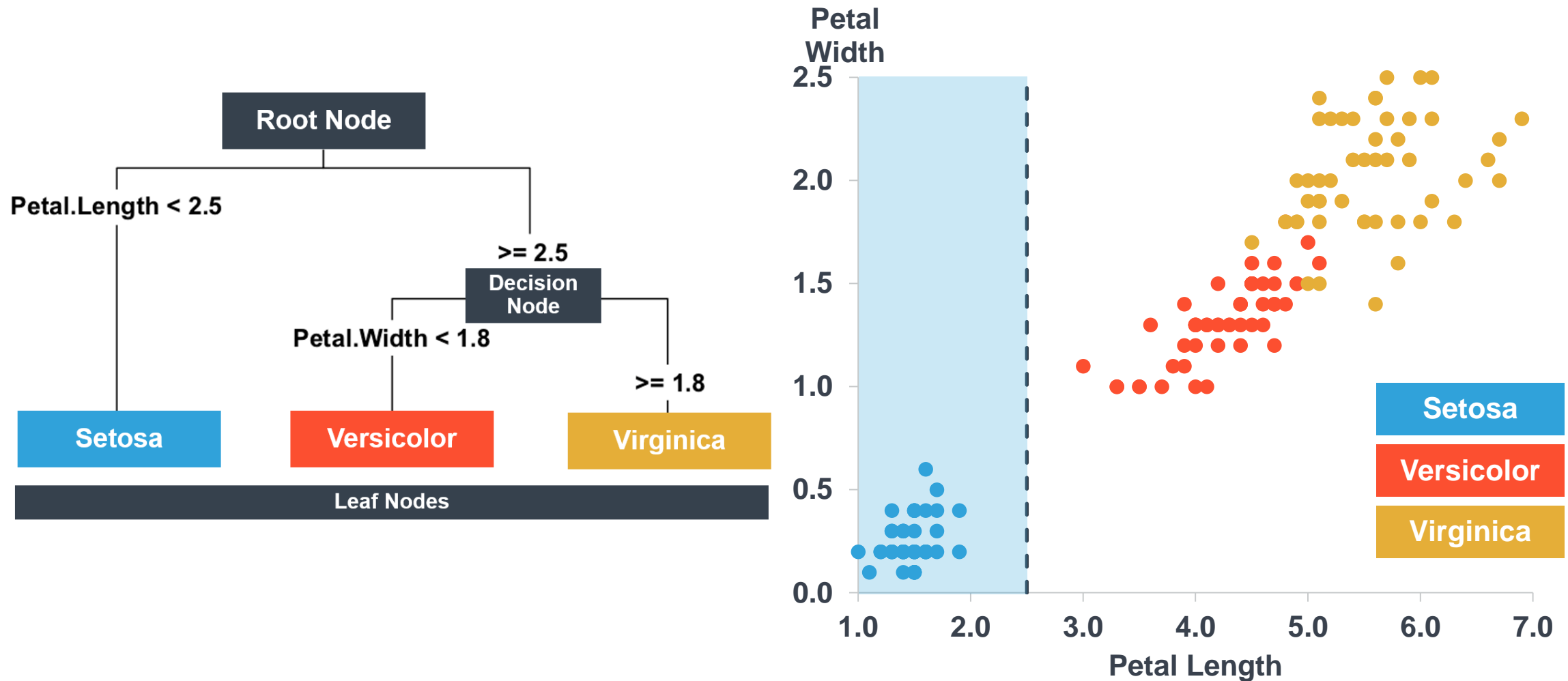
Decision trees provide another method for classification



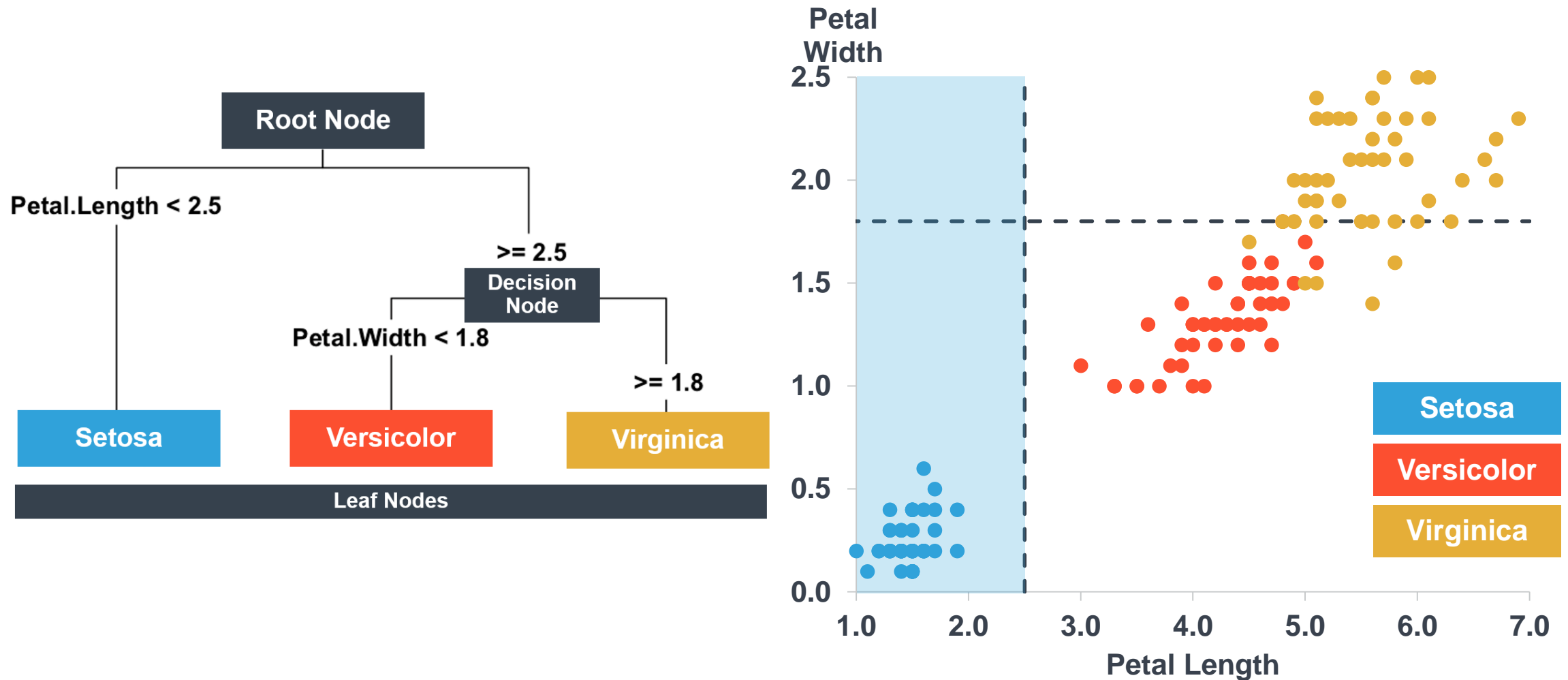
Decision trees provide another method for classification



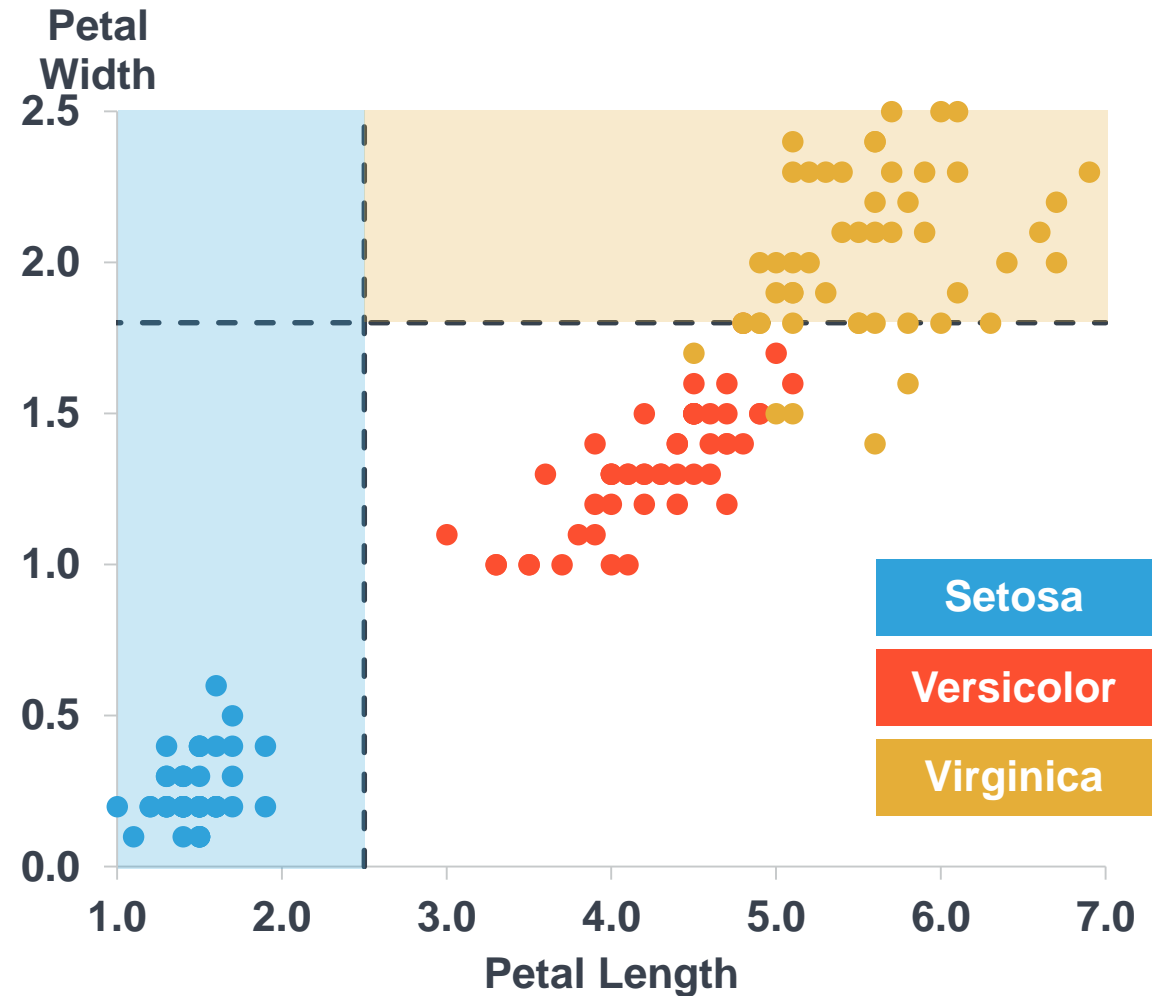
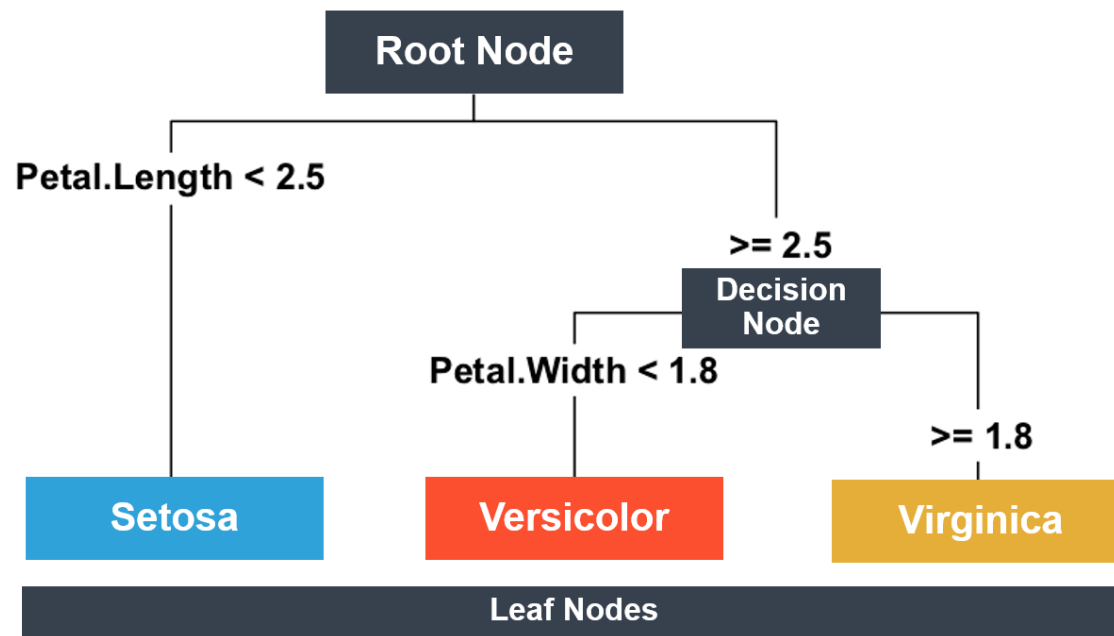
Decision trees provide another method for classification



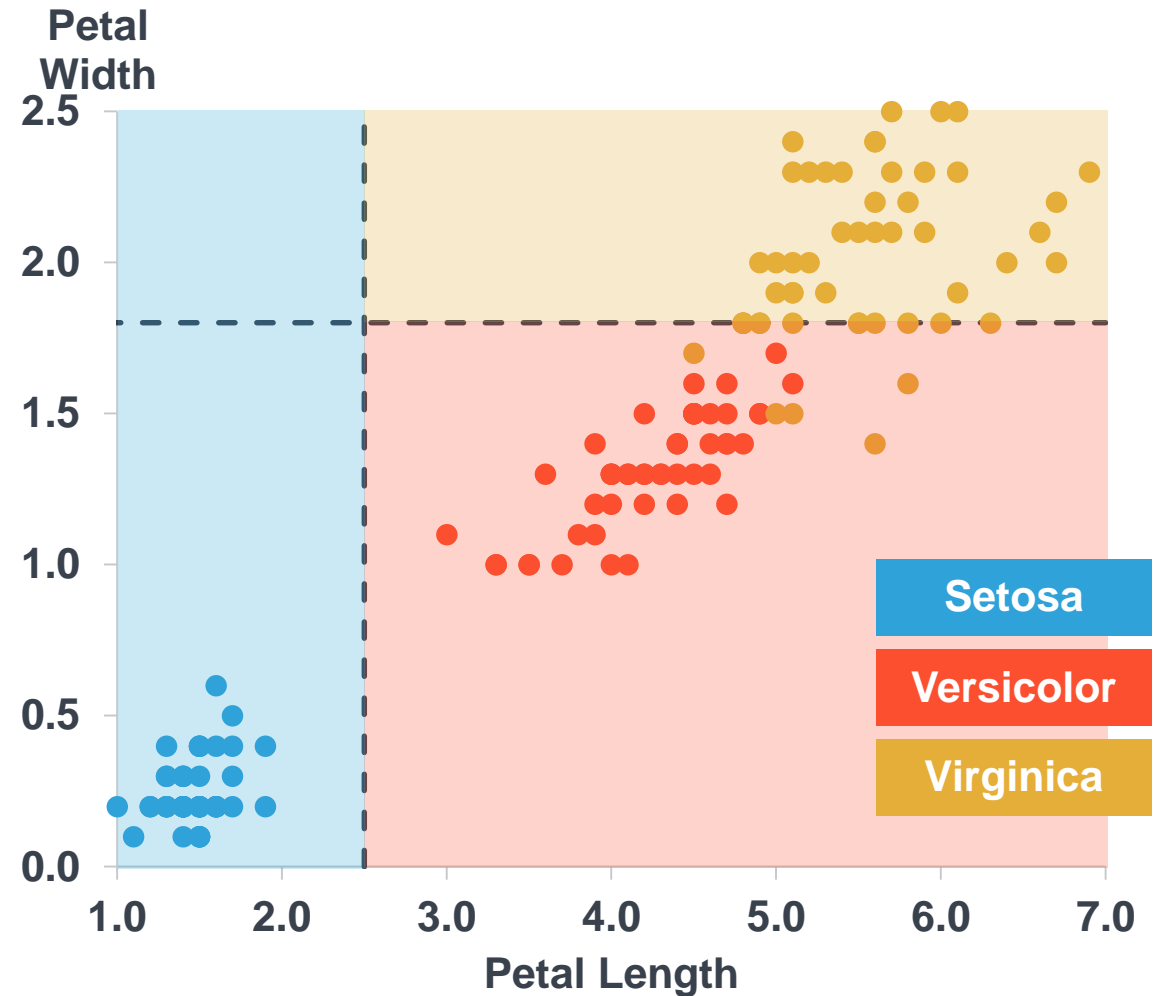
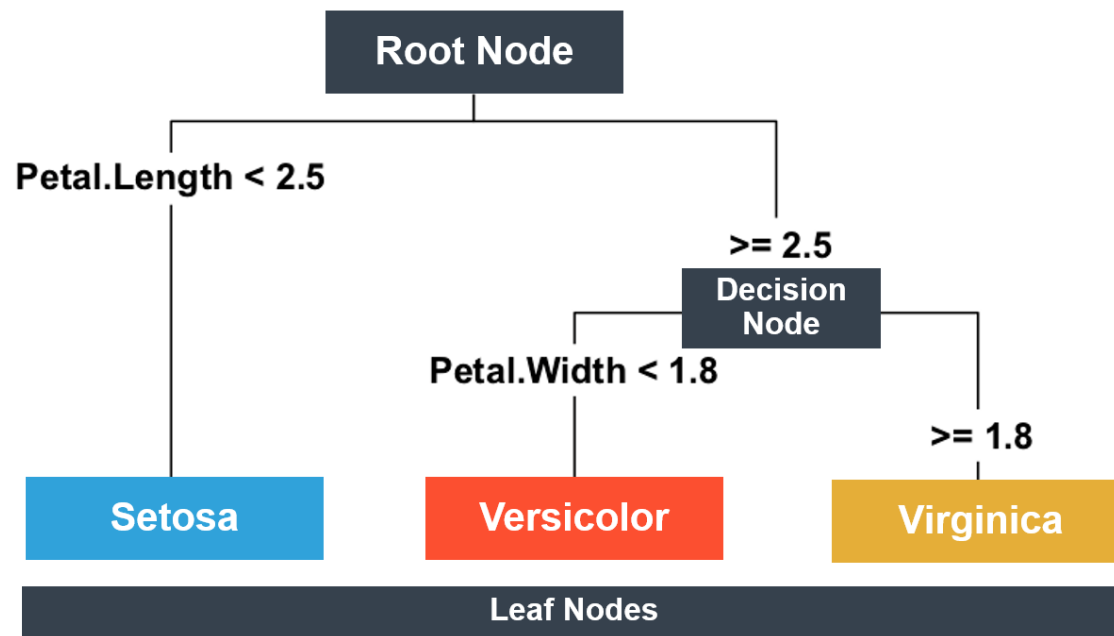
Decision trees provide another method for classification



Decision trees provide another method for classification

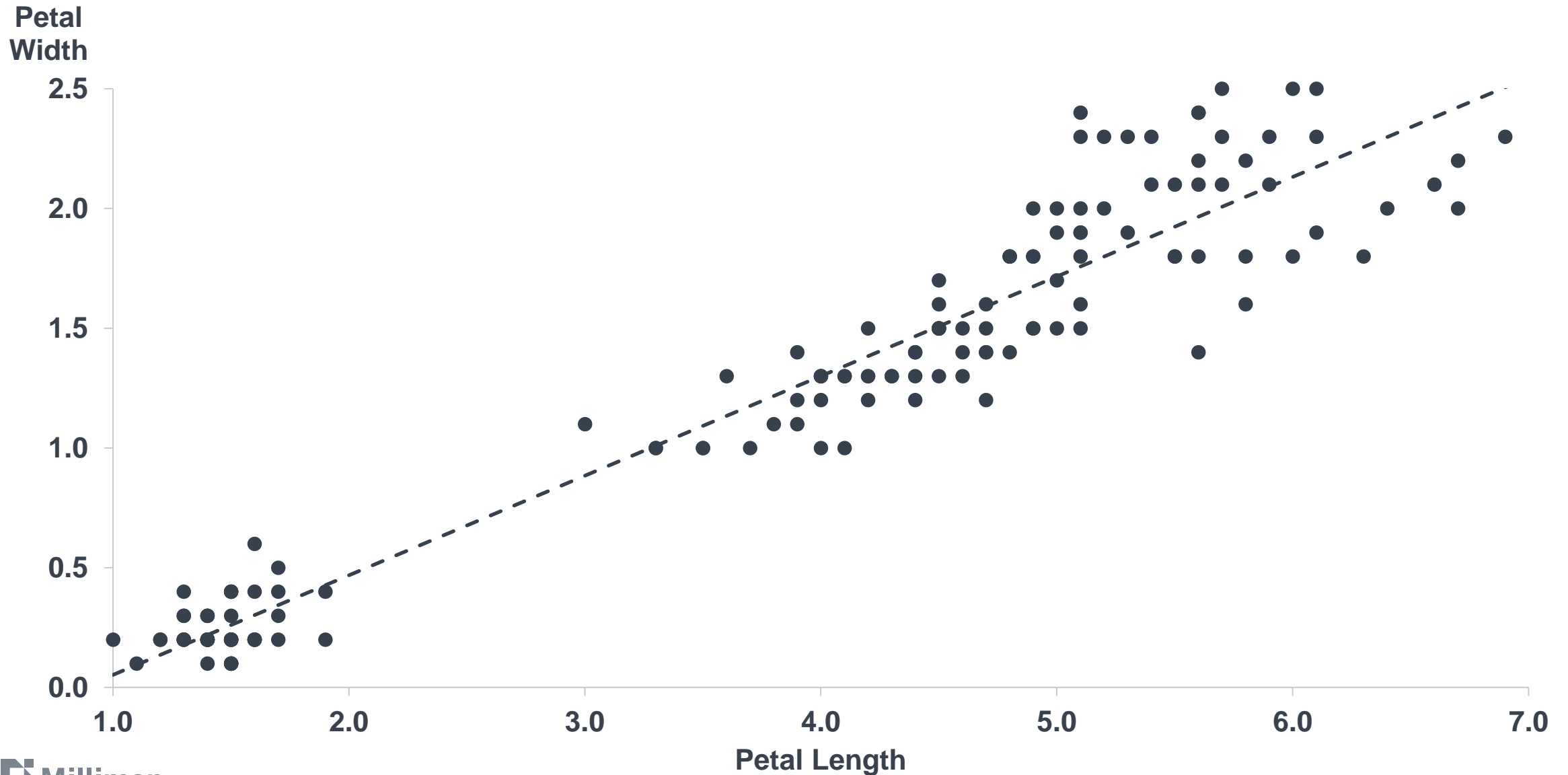


Decision trees provide another method for classification

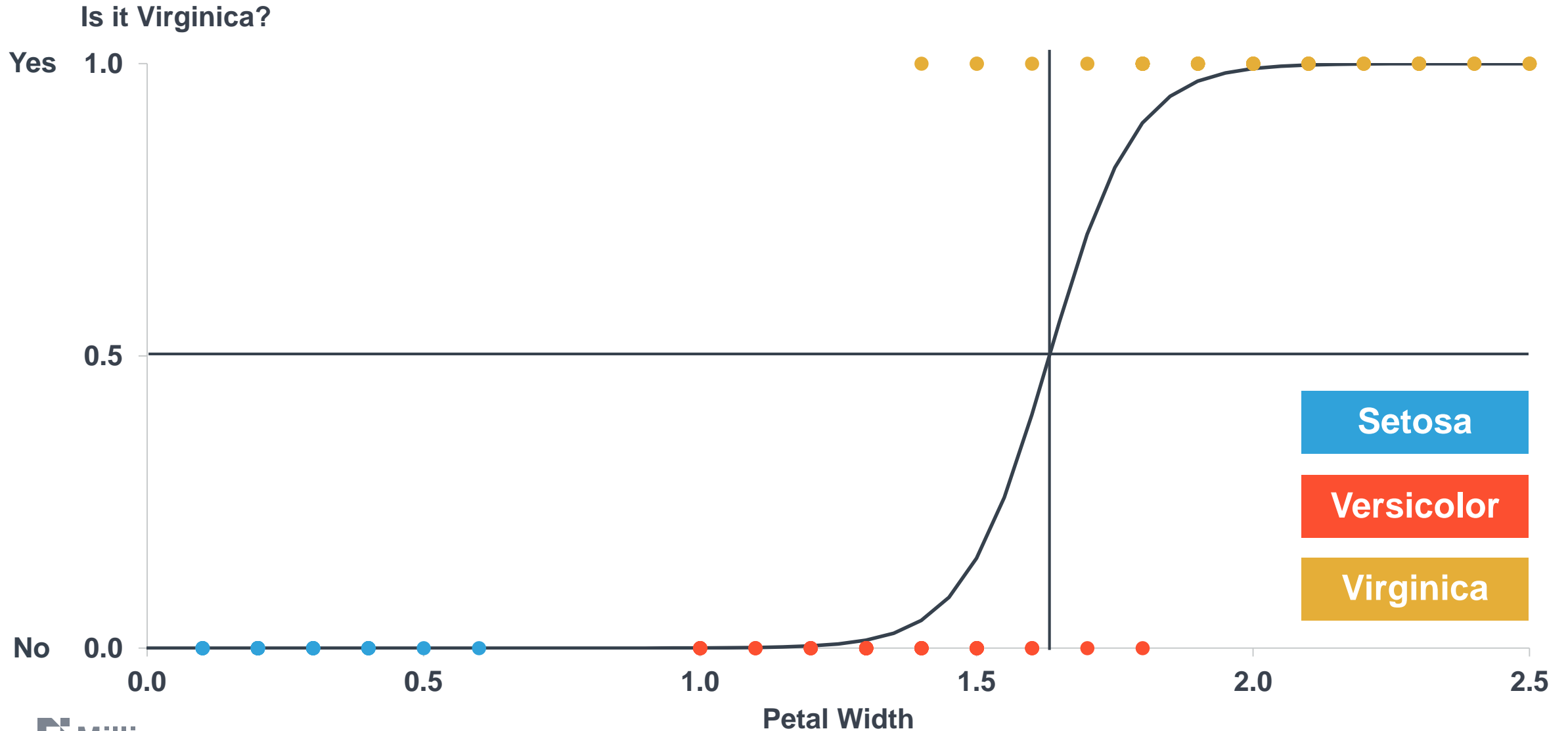


Logistic regression

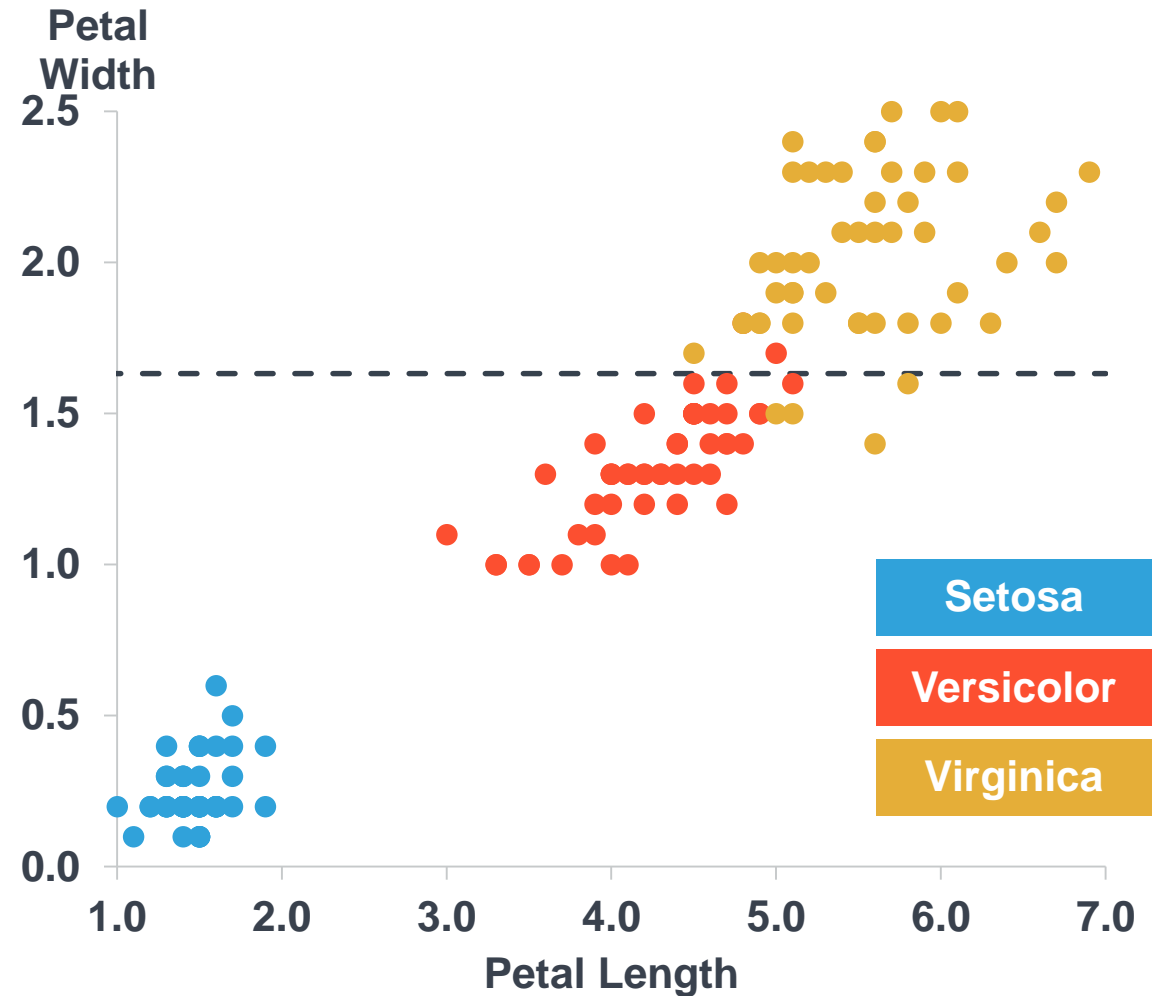
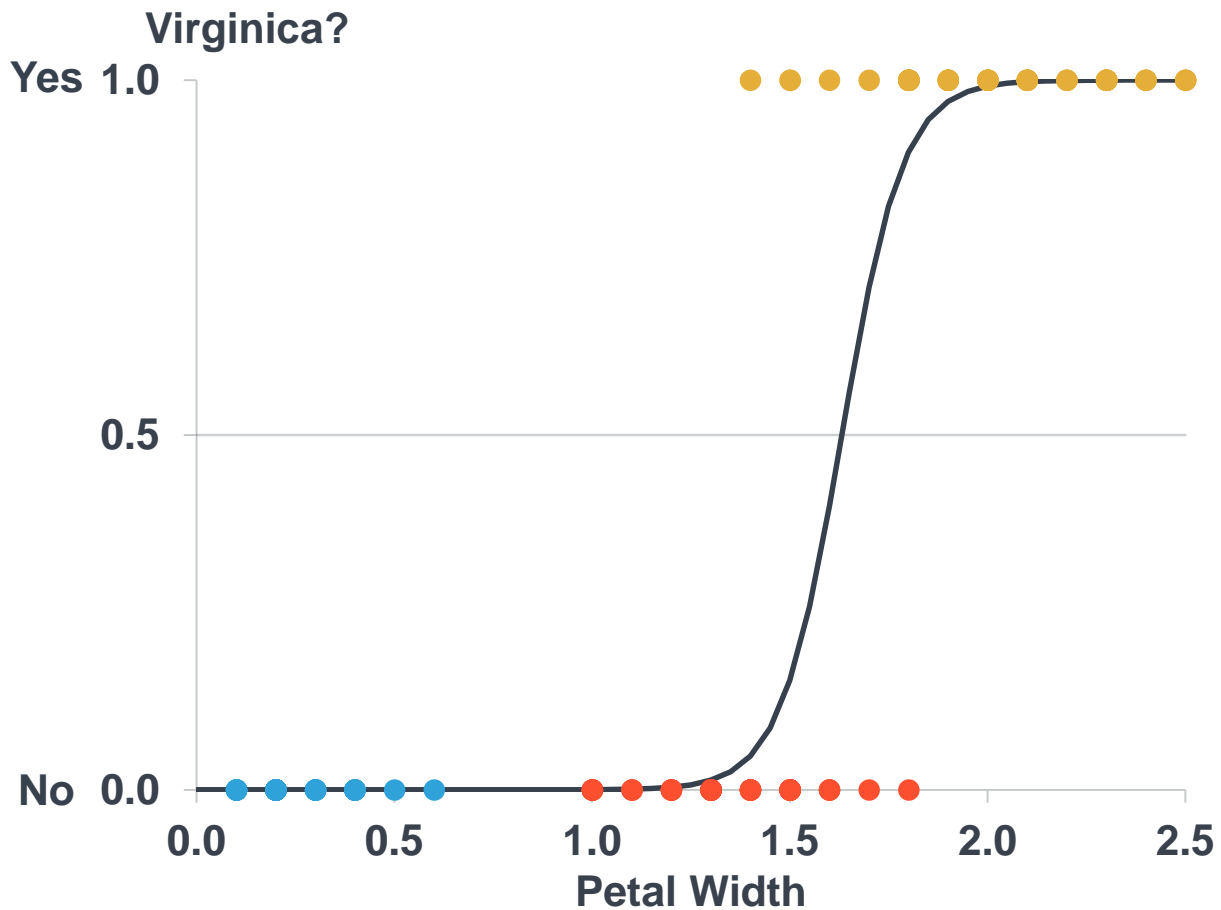
Linear regression uses line of best fit to make numerical predictions



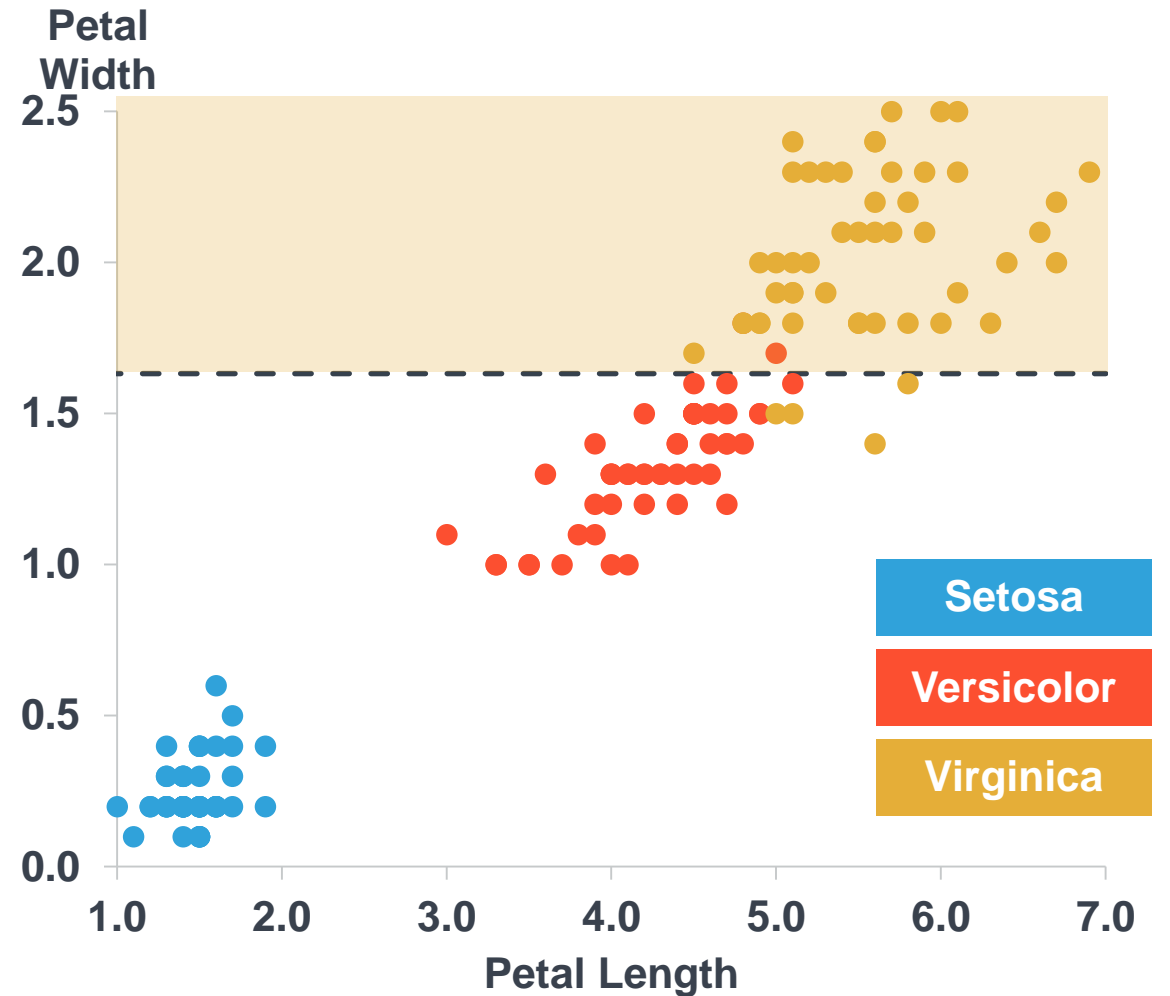
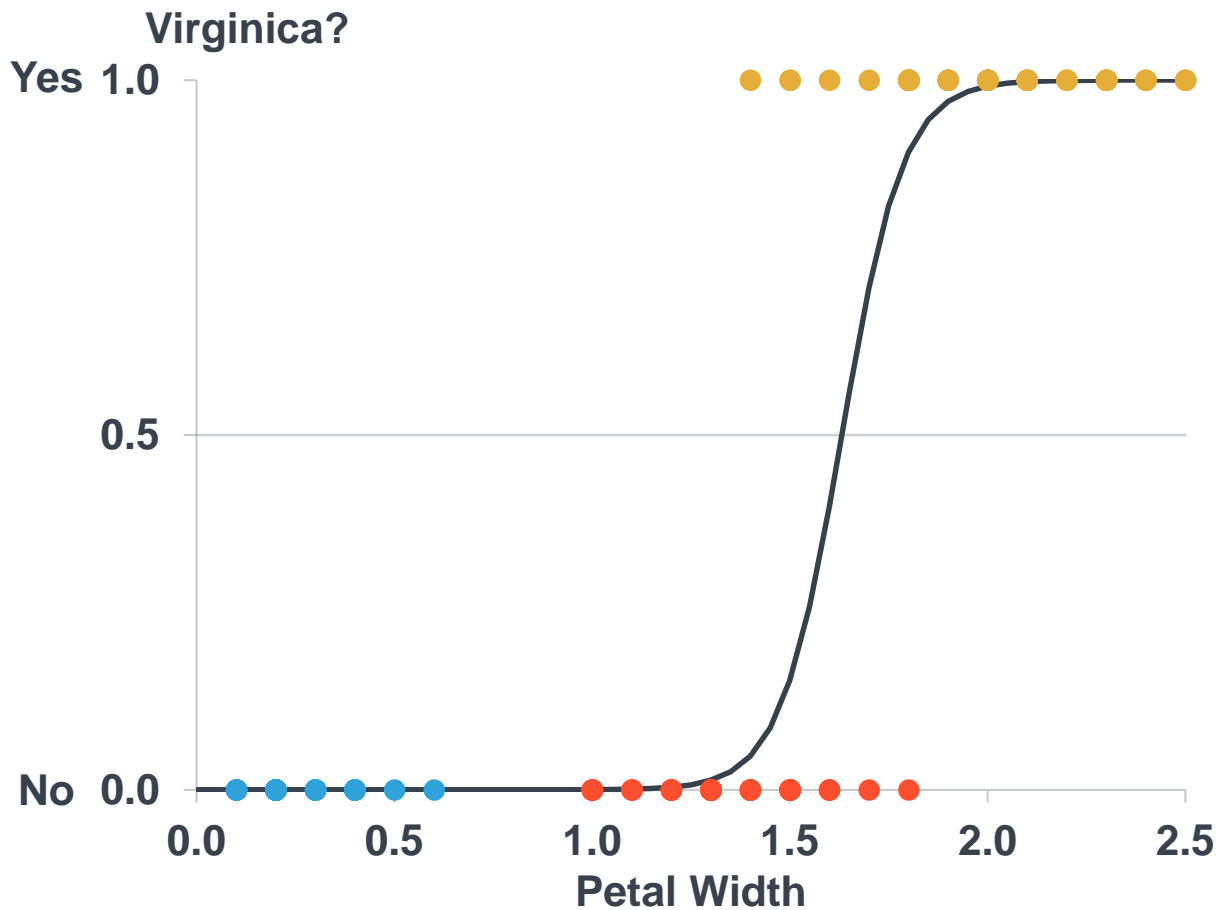
Logistic regression is a classification algorithm



Is it virginica?

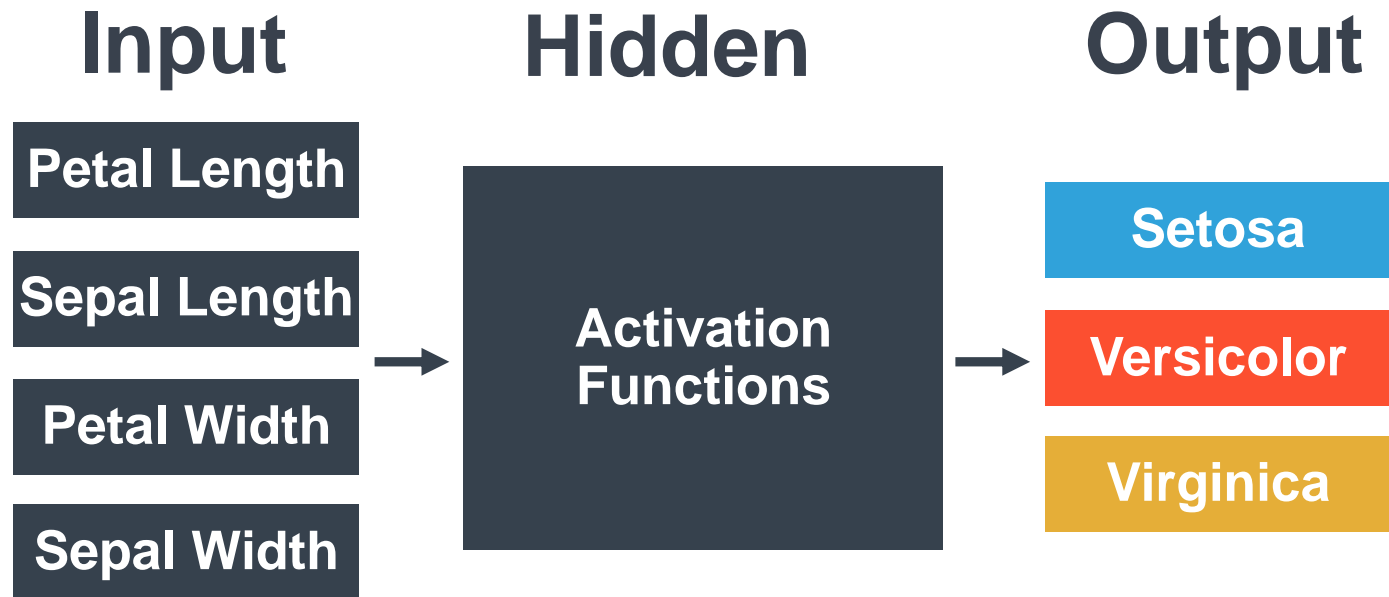


Is it virginica?

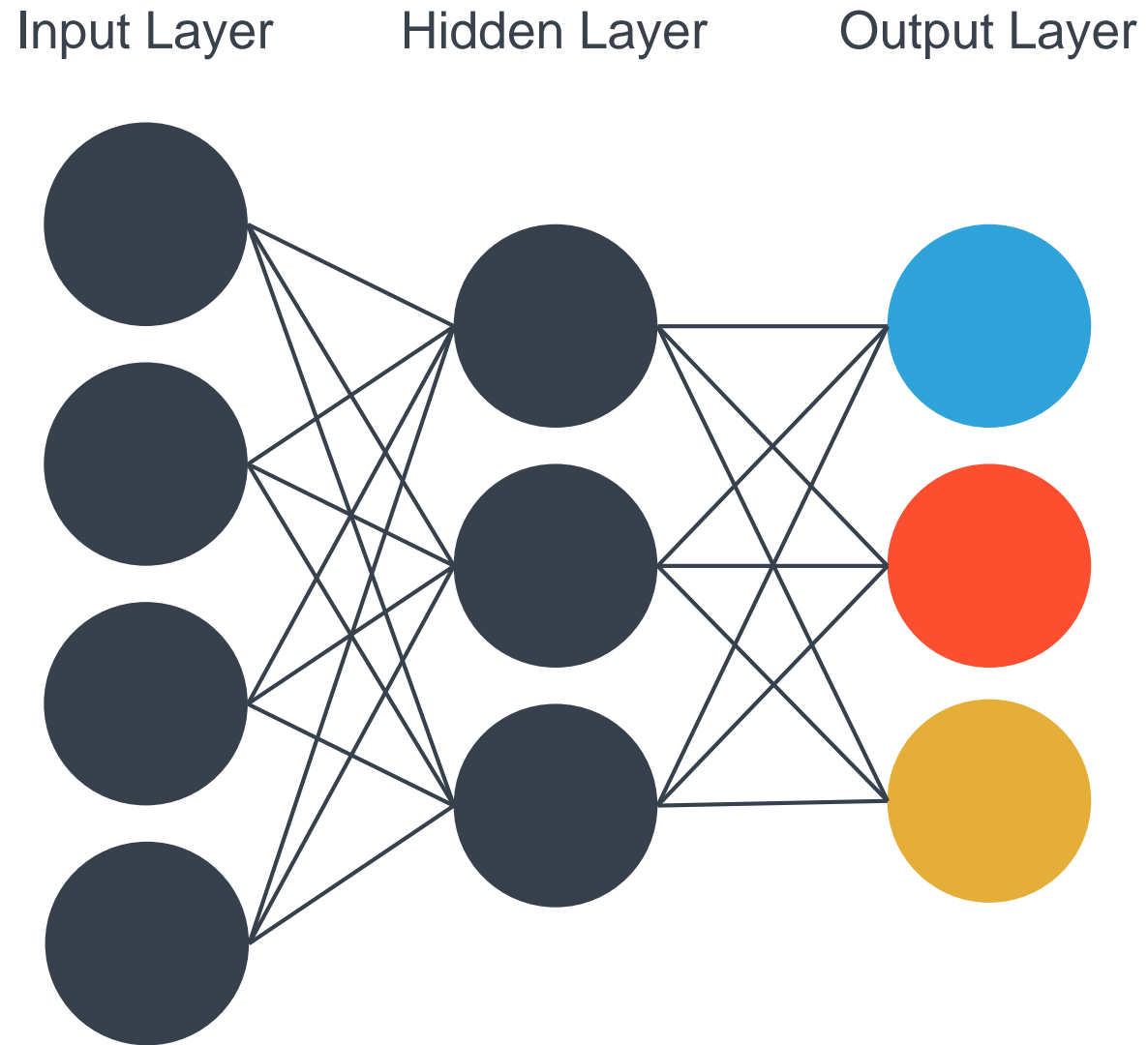


Neural networks

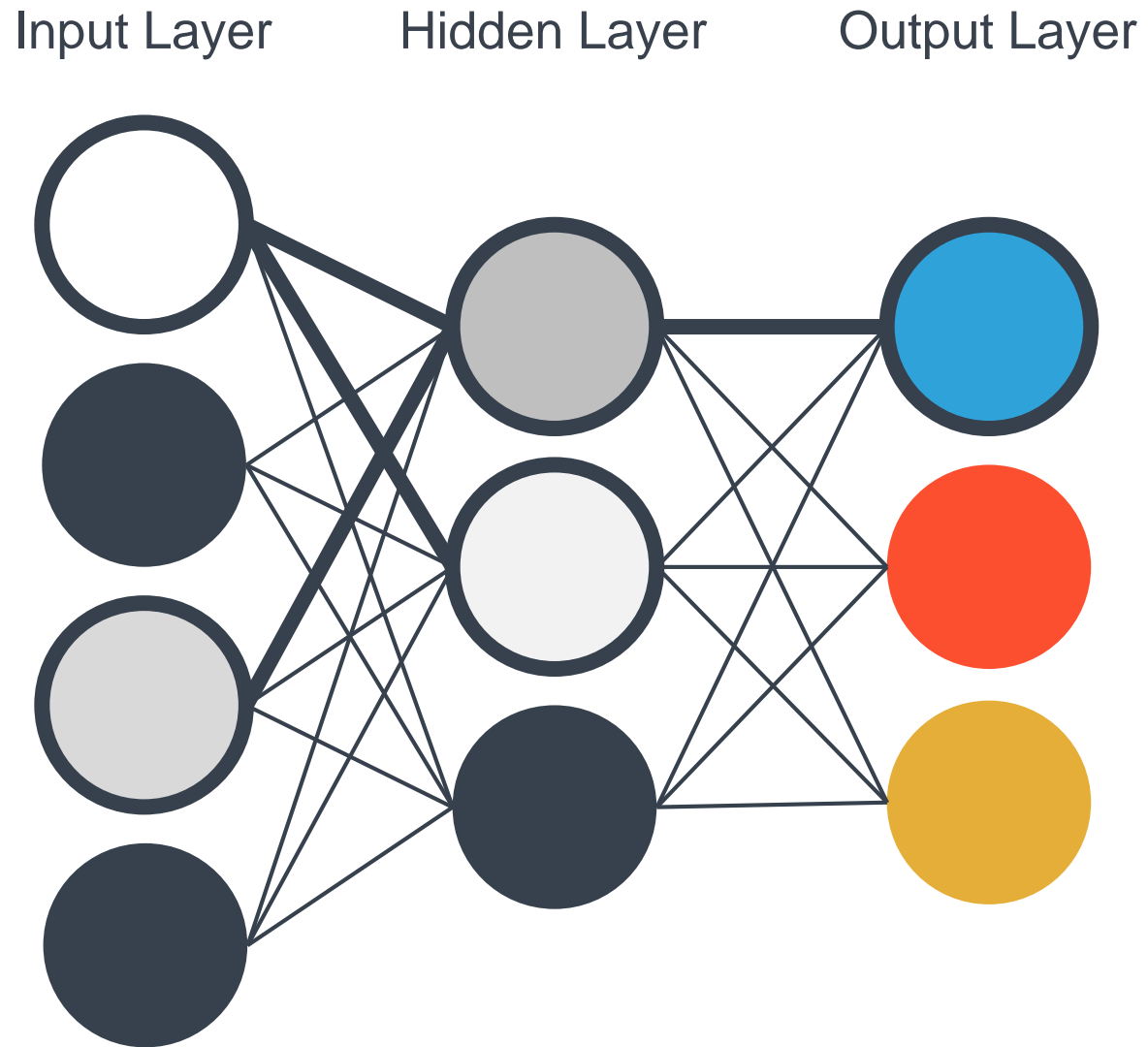
Neural can model non-linear situations



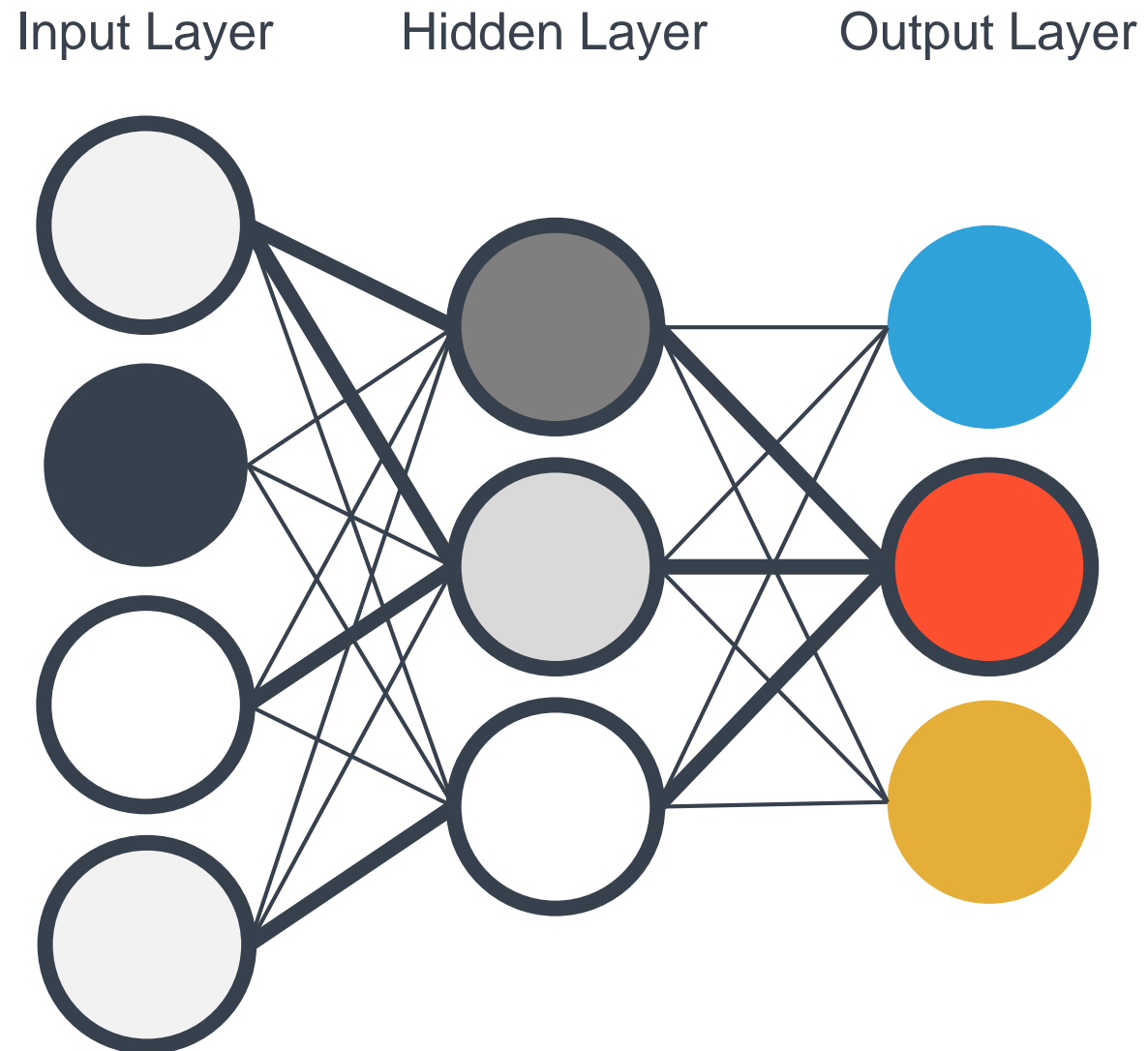
Neural Networks



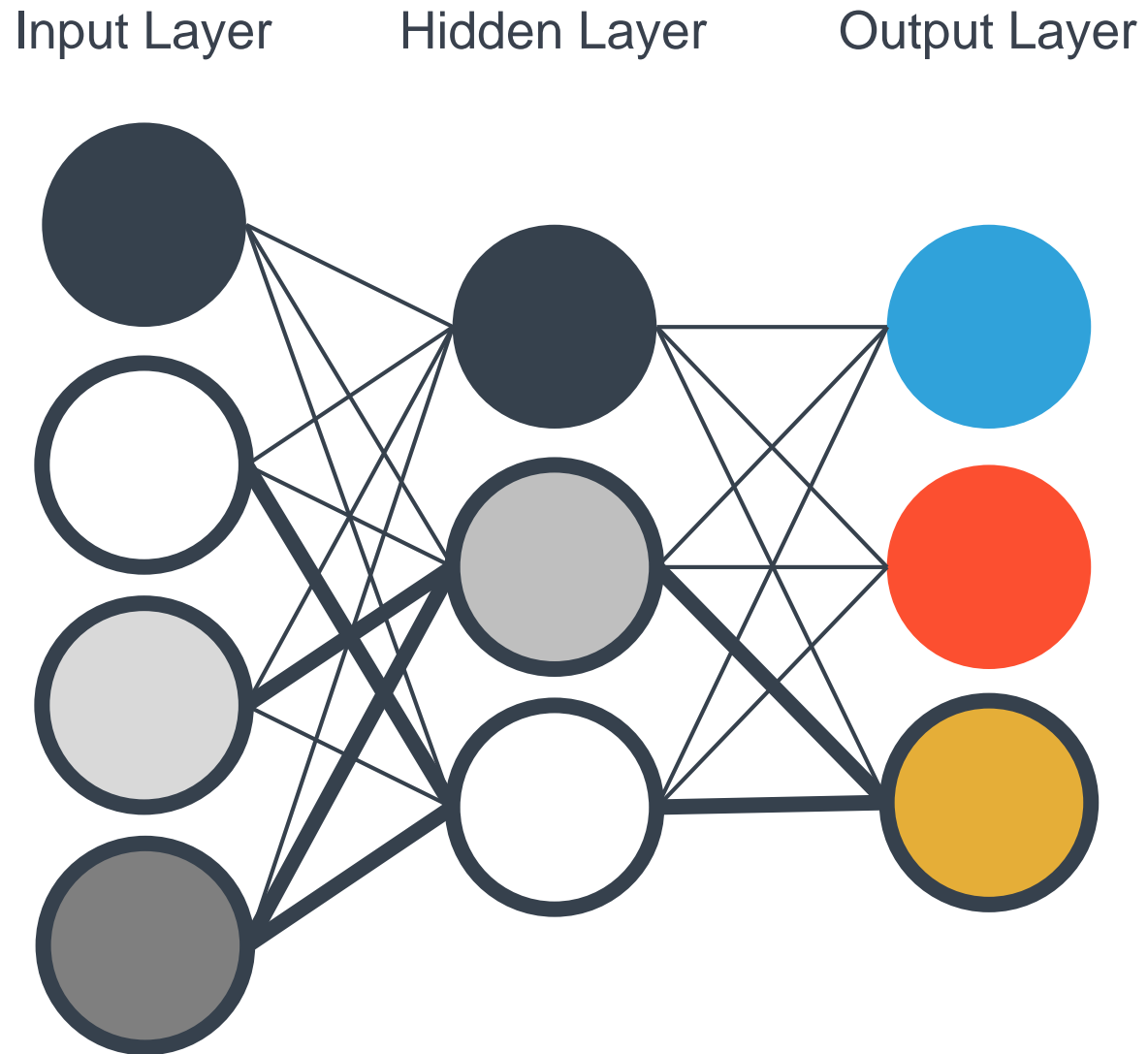
Neural Networks



Neural Networks



Neural Networks



Summary of methods

	Unsupervised	Deep	Regression
k-Nearest Neighbors			
k-Means Clustering	●		
Decision Trees			●
Linear Regression			●
Logistic Regression			
Neural Networks	●	●	●

Next steps

Possible machine learning applications in healthcare include

Predicting non-adherent drug event before it happens

Predict opioid drug abuse before it happens

Estimate member persistency based on member and dependent characteristics

Project medical costs using personal and clinical data

Develop clinical best-practices by linking clinical and financial data

