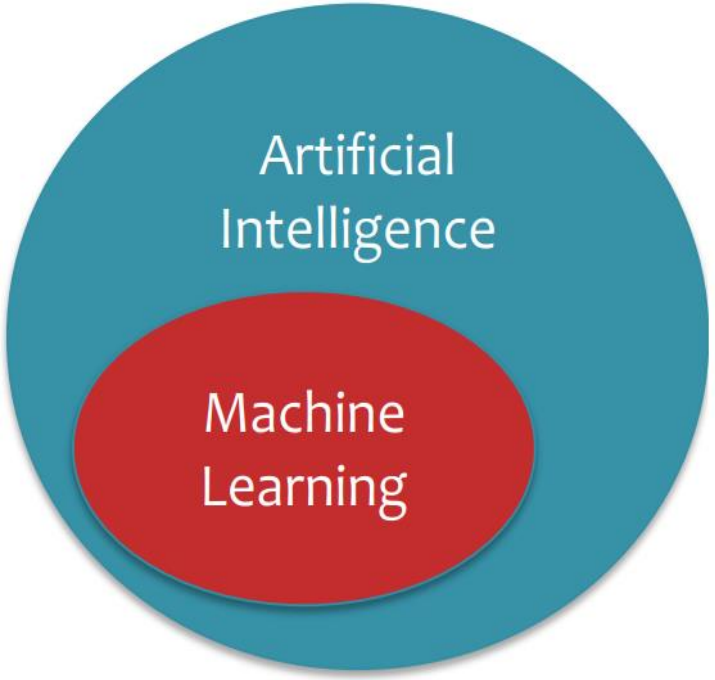
The Intuition Behind Machine Learning

Dr. Muhammad Wasim

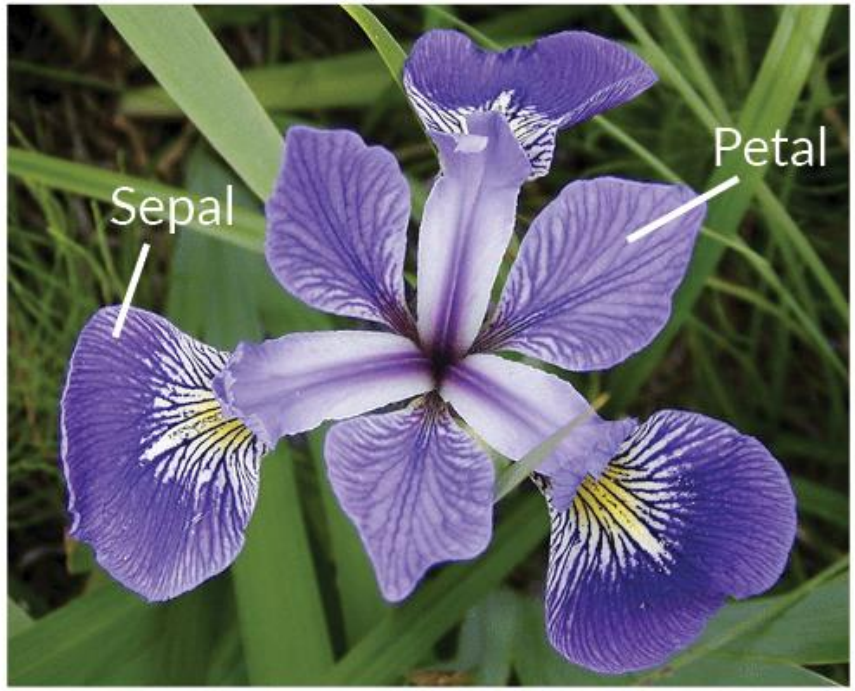
Artificial Intelligence

- The Basic goal of AI is to develop intelligent machines which consists of following goals:
  - Perception
  - Reasoning
  - Control / Motion / Manipulation
  - Planning
  - Communication
  - Creativity
  - Learning



Example: Iris Plant Categories

Can you identify different types of Iris flowers?



Iris Versicolor

Iris Setosa

Iris Virginica

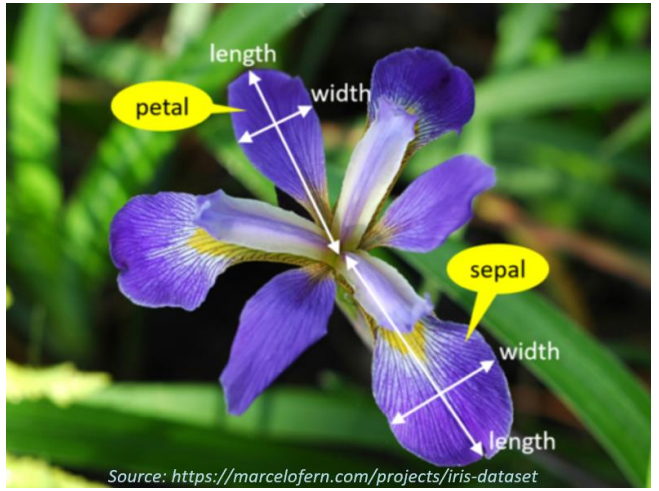
Iris Plant Categories Labeled Dataset

Features / Independent Variables / X

Class / Label / y

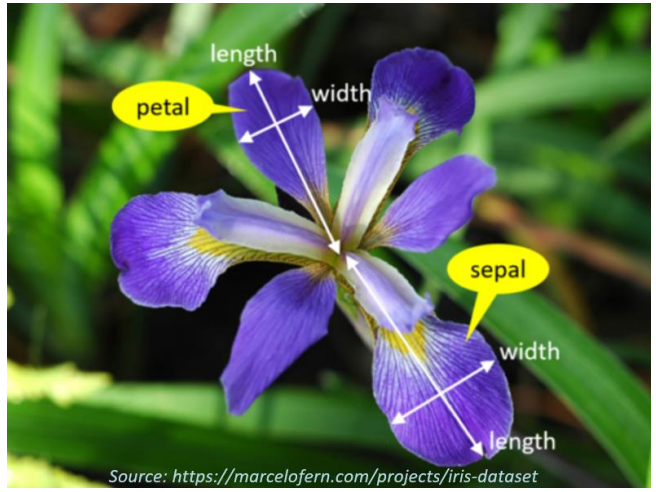
Sepal Length	Sepal Width	Petal Length	Petal Width	Type
5.1	3.5	1.4	0.2	Iris setosa
4.9	3.0	1.4	0.2	Iris setosa
7.0	3.2	4.7	1.4	Iris versicolor
6.4	3.2	4.5	1.5	Iris versicolor
6.3	3.3	6.0	2.5	Iris virginica
5.8	3.3	6.0	2.5	Iris viginica

Example / Instance



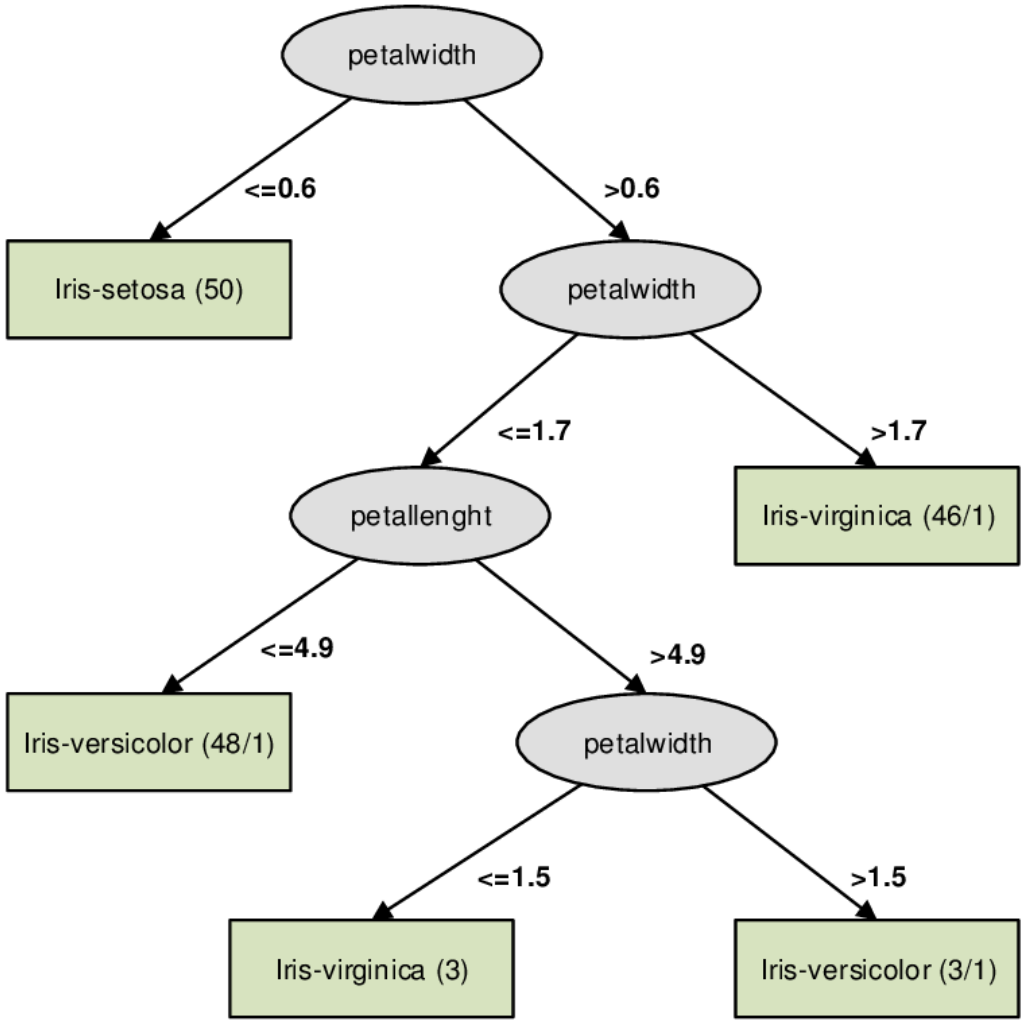
Iris Plant Categories

- R. A Fisher[1] collected the data of three categories of Iris plant.
- The dataset[2] contains the measurements of 50 flowers of each category (balanced dataset) so a total of 150 examples or instances.
- The dataset is multivariate and has four features (X) and each example is labeled with one of three classes (y).
- Now, can you classify the iris plants using this dataset?



1- R. A. Fisher (1936). "The use of multiple measurements in taxonomic problems"
2- https://archive.ics.uci.edu/ml/datasets/iris

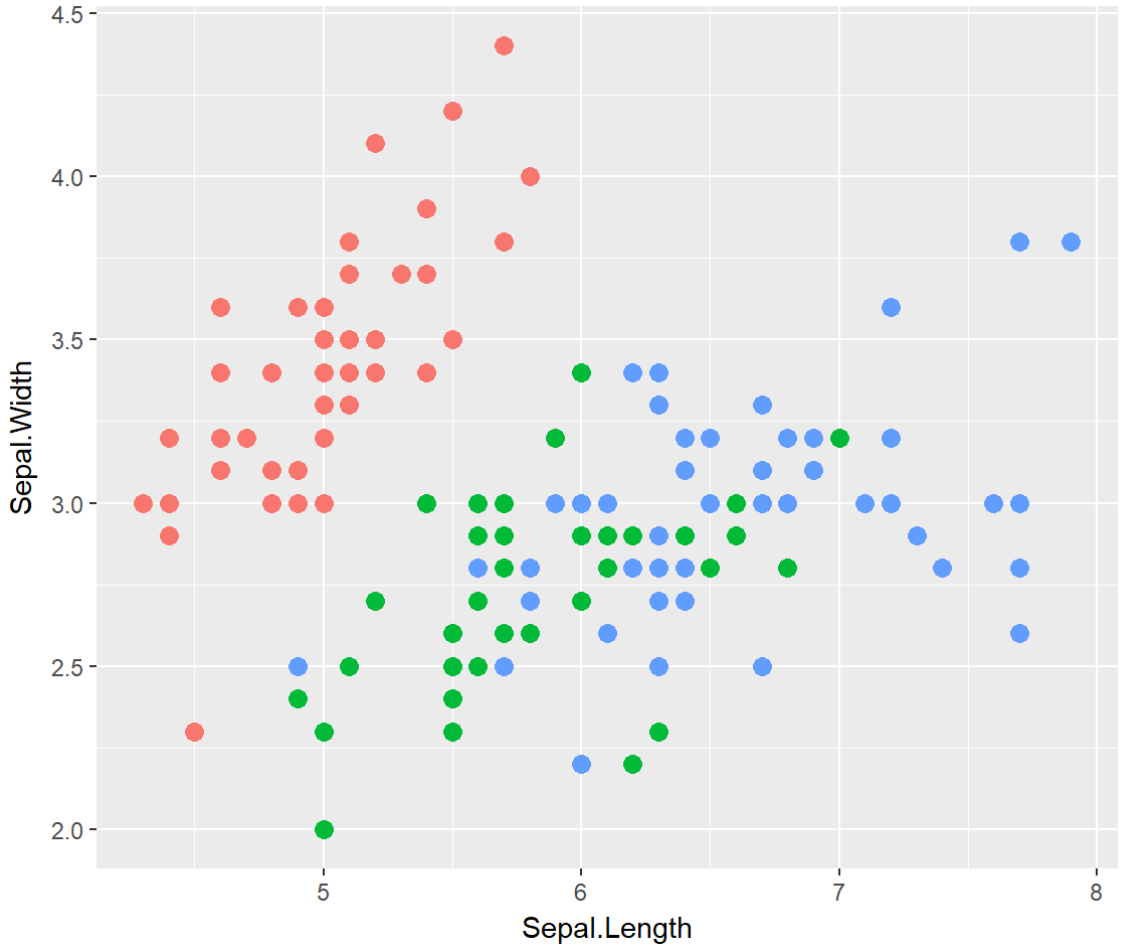
If-else conditions based on some threshold?

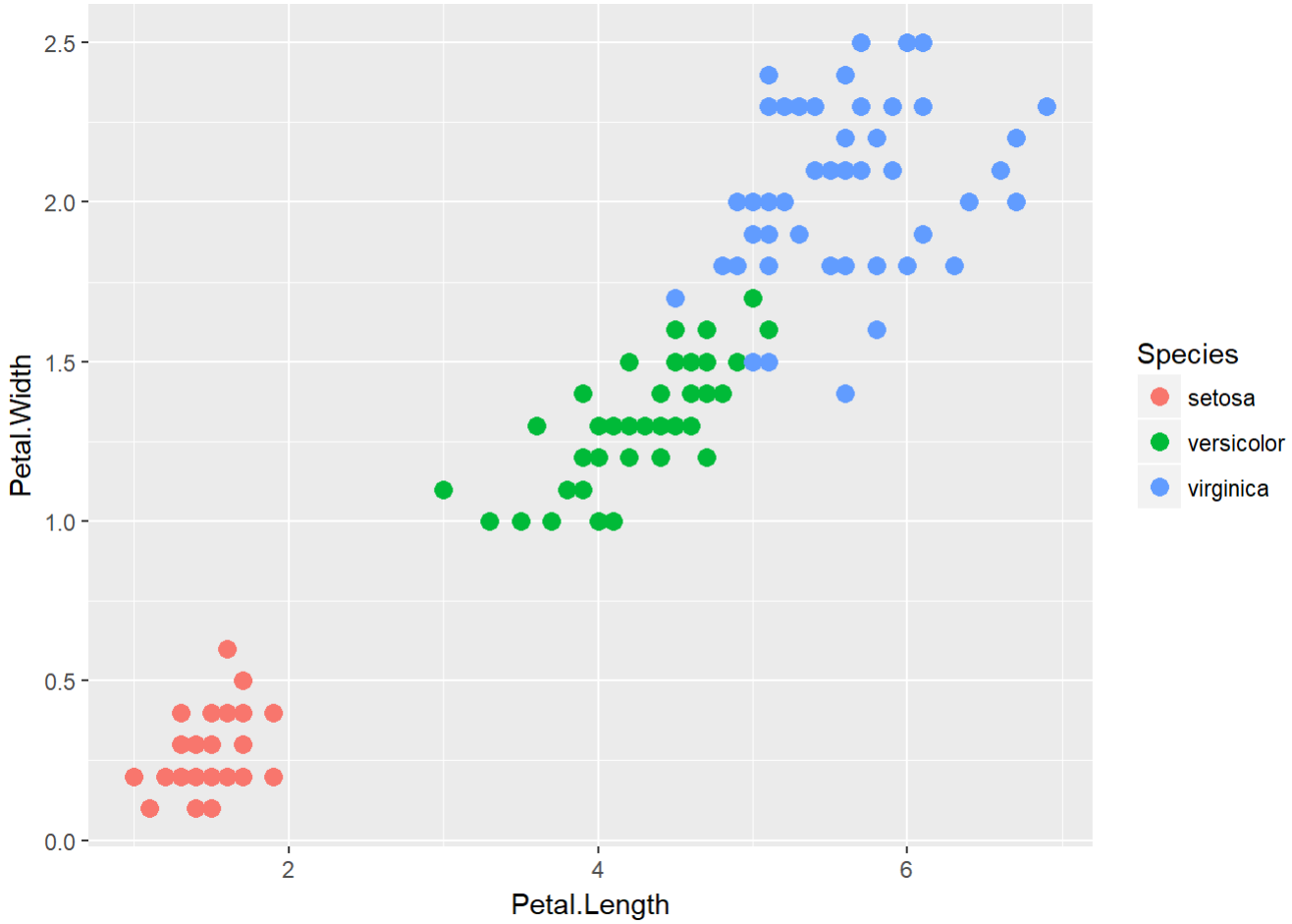


Decision Tree Model

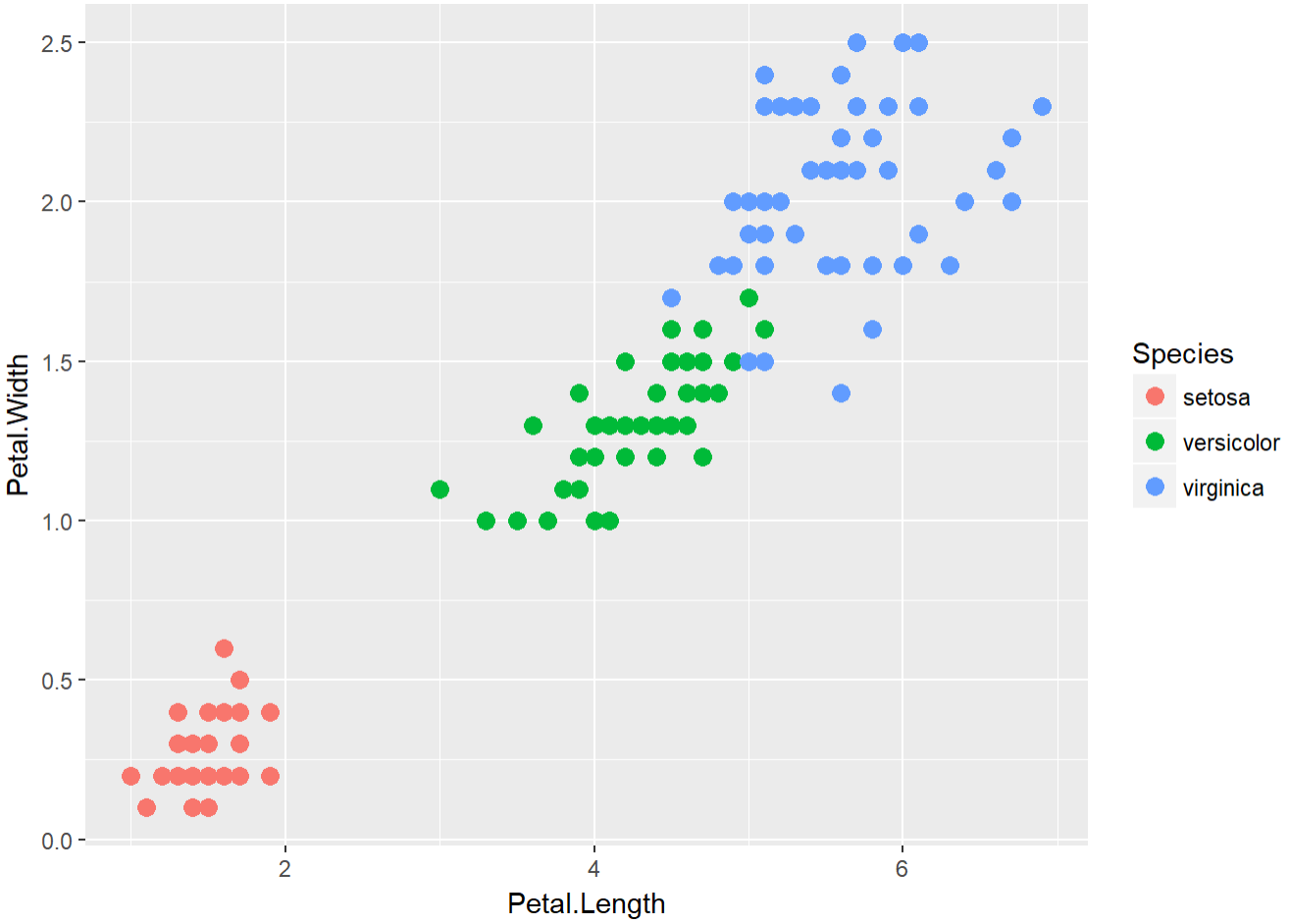
- We can write a program that encodes a set of rules that are useful to solve this problem.
- In many cases it is very difficult to specify those rules
- Learning systems are not directly programmed to solve a problem,
  instead such systems develop their own program based on the examples of how they should behave.
- Want to implement **unknown function**, only have access e.g., to sample input-output pairs (training examples)
- Learning simply means incorporating information from the training examples into the system.

Is there any other way we can learn the patterns in data?





Is there any other way we can learn the patterns in data?

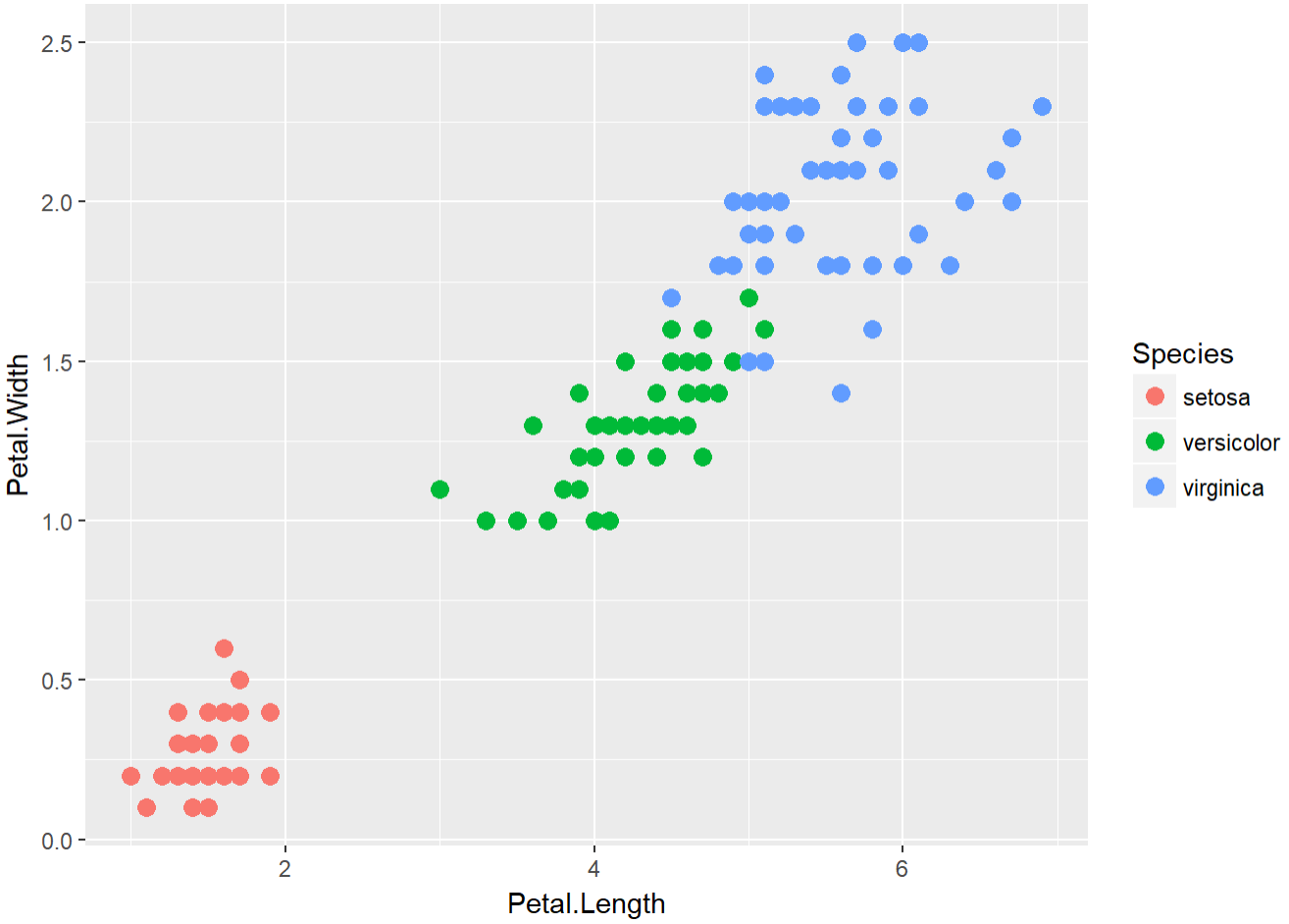


- OK, we have three classes. In case, we had two classes (binary classification), we could identify the decision boundary between the two classes using a line.
- If we identify the y-intercept (c) and slope (m) of the line which separates the two classes, we can easily classify the new example in test data.

$$y = mx + c$$

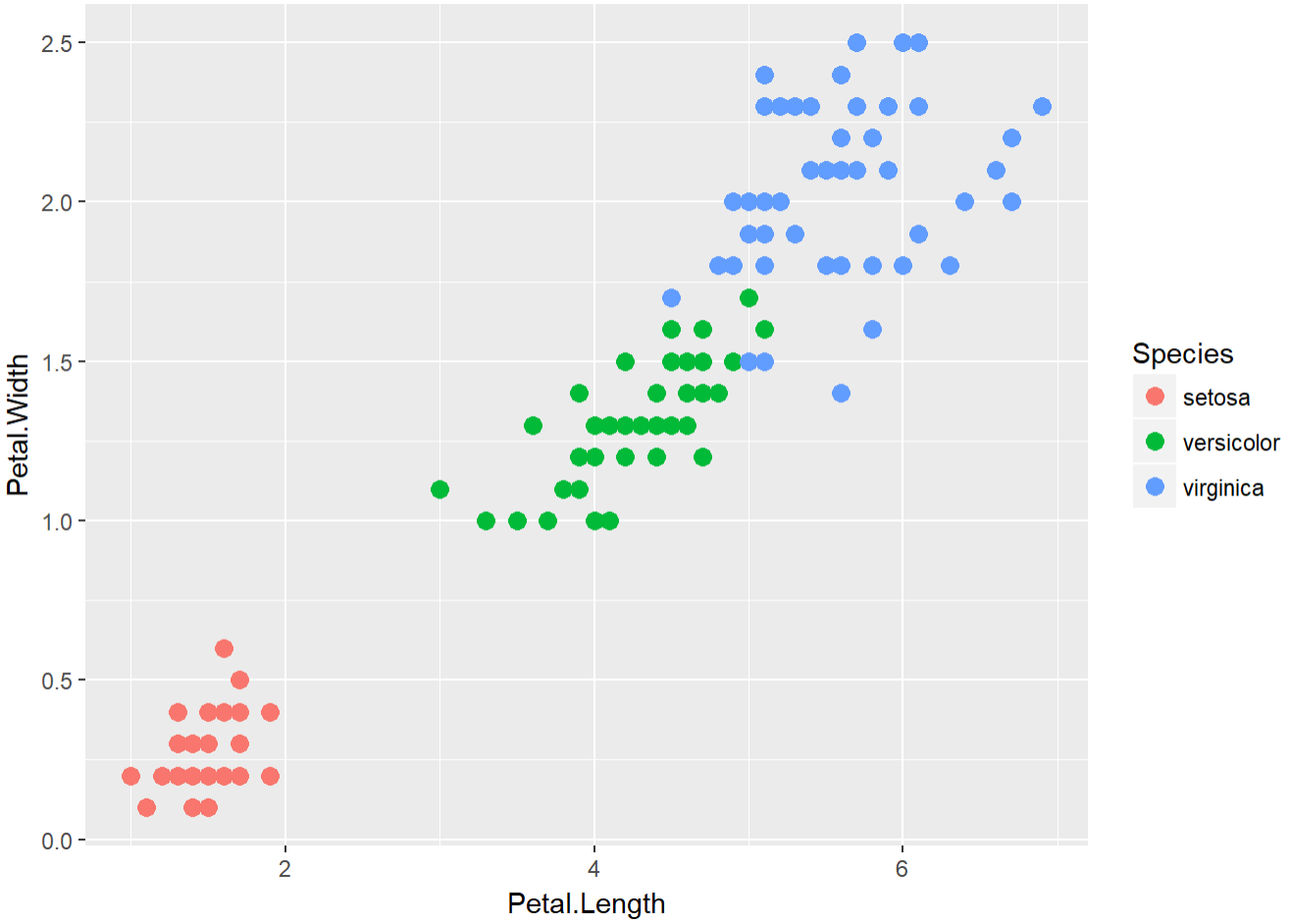
- Linear and Logistic Regression

Is there any other way we can learn the patterns in data?



- So, Will any line sufficient to separate the two different type of data?
- What will be the ideal line to sperate the data?
- Support Vector Machine (SVM)

Is there any other way we can learn the patterns in data?



- Duck test
- If it looks like a duck, swims like a duck, and quacks like a duck, then it probably is a duck.
- If we can measure the distance of the unseen example with the examples in the training data, we can check what it looks like.
- K-nearest neighbors

Is there any other way we can learn the patterns in data?

x1	x2	x3	x4	y
5.1	3.5	1.4	0.2	Iris setosa
4.9	3.0	1.4	0.2	Iris setosa
7.0	3.2	4.7	1.4	Iris versicolor
6.4	3.2	4.5	1.5	Iris versicolor
6.3	3.3	6.0	2.5	Iris virginica
5.8	3.3	6.0	2.5	Iris viginica

- We can also calculate the probability of class given the features.
- That is we can calculate three probabilities and select the class with highest probability:

$$P(IrisSetosa|x_1, x_2, x_3, x_4)$$
$$P(IrisVersicolor|x_1, x_2, x_3, x_4)$$
$$P(IrisVirginica|x_1, x_2, x_3, x_4)$$

- Naïve Bayes Classifier

Is there any other way we can learn the patterns in data?

x1	x2	x3	x4	y
5.1	3.5	1.4	0.2	Iris setosa
4.9	3.0	1.4	0.2	Iris setosa
7.0	3.2	4.7	1.4	Iris versicolor
6.4	3.2	4.5	1.5	Iris versicolor
6.3	3.3	6.0	2.5	Iris virginica
5.8	3.3	6.0	2.5	Iris viginica

x5	x6	y
0.11	0.22	Iris setosa
0.8	0.33	Iris setosa
0.1	0.88	Iris versicolor
0.08	0.75	Iris versicolor
0.9	0.52	Iris virginica
0.85	0.62	Iris viginica

- We can also learn feature representation (derive new features) from existing features
- These new features can then be used for classification.

- Perceptron and Neural Networks