**WIKI SUPERVISED LEARNING DEFINITION**

* **TRANING DATA (historical/lean)**
* **test data (for predict)**
* **supervised**
* **unsuperviseD**

Supervised learning is the Data mining task of inferring a function from **labeled training data**.

The training data (historical/learn data) consist of a set of training examples.

In supervised learning, each example is a pair consisting of an input object (typically a vector/Column) and a desired output value (also called the **supervisory signal**).

A **supervised learning algorithm** analyzes the training data and produces an inferred function, which can be used for mapping new examples. An optimal scenario will allow for the algorithm to correctly determine the class labels for**unseen instances**. This requires the learning algorithm to generalize from the training data to unseen situations in a “reasonable” way.

**WIKI UNSUPERVISED LEARNING DEFINITION**

In Data mining, the problem of **unsupervised learning** is that of trying to find hidden structure in unlabeled data. Since the examples given to the learner are unlabeled, there is no error or reward signal to evaluate a potential solution.

**LET’S LEARN SUPERVISED AND UNSUPERVISED LEARNING WITH A REAL LIFE EXAMPLE**

[](https://dataaspirant.files.wordpress.com/2014/09/basket.jpg)

* + suppose you had a basket and it is fulled with some different kinds of fruits, your task is to arrange them as groups.
  + For understanding let me clear the names of the fruits in our basket.
  + We have four types of fruits. They are:**apple, banana, grape and cherry.**

**SUPERVISED LEARNING :**

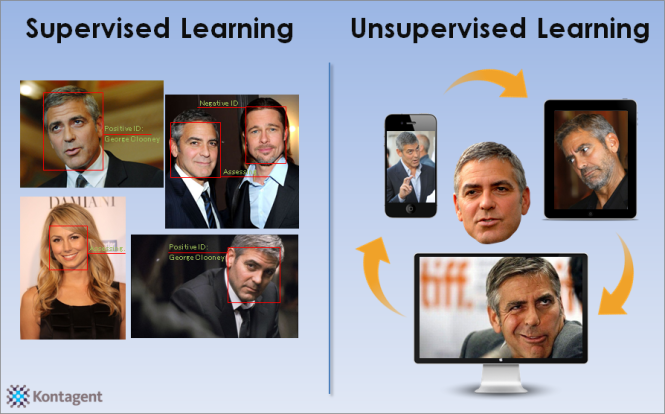
* You already learn from your previous work about the physical characters of fruits.
* So arranging  the same type of fruits at one place is easy now.
* Your previous work is called as **training data** in data mining.
* so you already learn the things from your train data, this is because of **response variable.**
* **Y(response/dependent)~x(input/independent)**
* Response variable mean just a **decision variable.**
* You can observe response variable below (**FRUIT NAME**) .

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **NO.** | **SIZE** | **COLOR** | **SHAPE** | **FRUIT NAME** |
| 1 | Big | Red | Rounded shape with a depression at the top | Apple |
| 2 | Small | Red | Heart-shaped to nearly globular | Cherry |
| 3 | Big | Green | Long curving cylinder | Banana |
| 4 | Small | Green | Round to oval,Bunch shape Cylindrical | Grape |

* Suppose you have  taken an new fruit from the basket then you will see the size , color and shape of that particular fruit.
* If  size  is Big , color is Red , shape is rounded shape with a depression at the top, you will conform the fruit name as apple and you will put in apple group.
* Likewise for other fruits also.
* Job of groping fruits was done and happy ending.
* You can observe in the table that  a column was labeled as “**FRUIT NAME**” this is called as response variable.
* If you learn the thing before from training data and then applying that knowledge to the test data(for new fruit), This type of learning is called as**Supervised Learning**.
* **Classification** come under Supervised learning.

**UNSUPERVISED LEARNING**

* Suppose you had a basket and it is fulled with some different types fruits, your task is to arrange them as groups.
* This time you don’t know any thing about that fruits, honestly saying this is the first time you have seen them.
* so how will you arrange them.
* What will you do first???
* You will take a fruit and you will arrange them by considering physical character of that particular fruit. suppose you have considered color.
* Then you will arrange them on considering base condition as **color.**
* Then the groups will be some thing like this.
* RED COLOR GROUP: apples & cherry fruits.
* GREEN COLOR GROUP: bananas & grapes.
* so now you will take another physical character such as **size** .
* RED COLOR AND BIG SIZE: apple.
* RED COLOR AND SMALL SIZE: cherry fruits.
* GREEN COLOR AND BIG SIZE: bananas.
* GREEN COLOR AND SMALL SIZE: grapes.
* job done happy ending.
* Here you didn’t know learn any thing before ,means no train data and no response variable.
* This type of learning is know unsupervised learning.
* clustering comes under unsupervised learning.

[](https://dataaspirant.files.wordpress.com/2014/09/george-clooney5.png)