Machine Learning Tutorial

Akshat Sinha
Gautam Gupta

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Activity Recognition chain

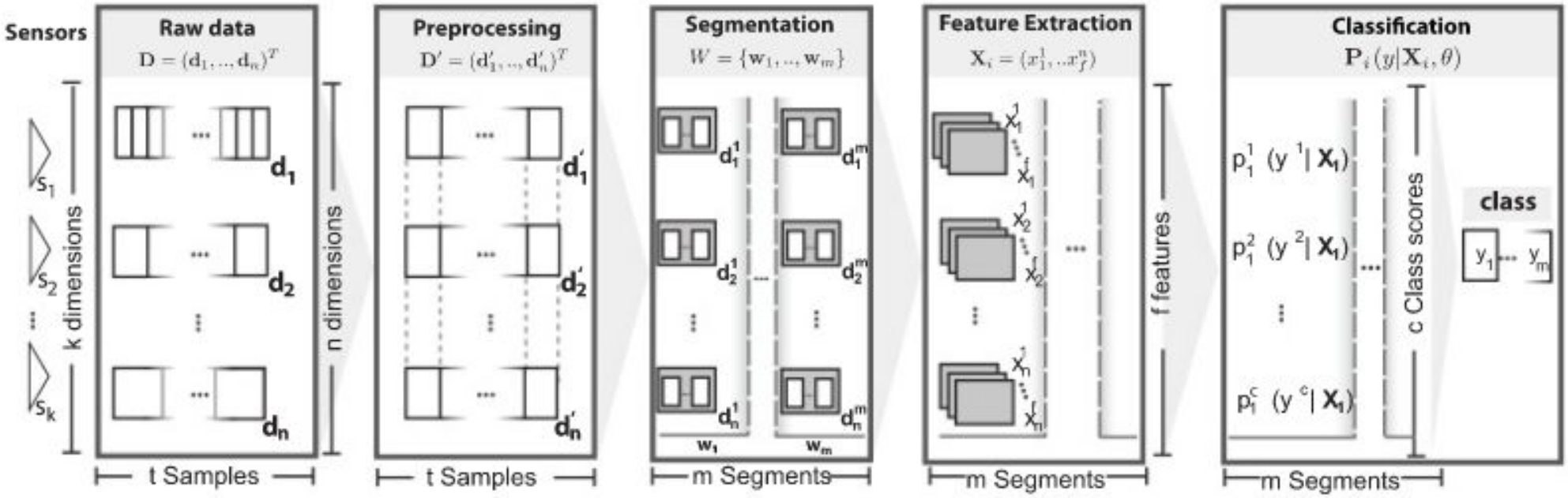


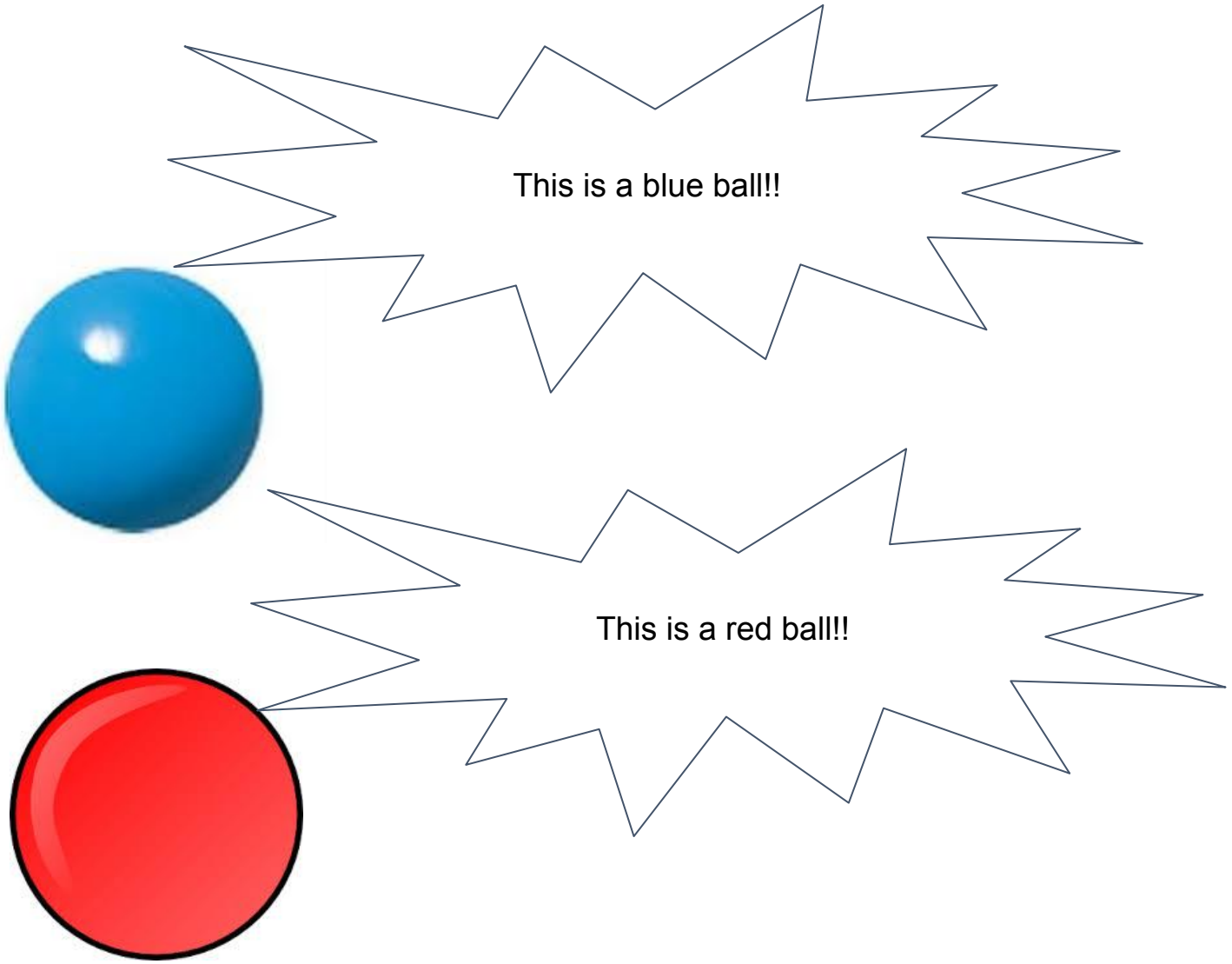
Fig. 1. Typical Activity Recognition Chain (ARC) to recognize activities from wearable sensors. An ARC comprises stages for data acquisition, signal preprocessing and segmentation, feature extraction and selection, training, and classification. Raw signals ($\mathbf{D}$) are first processed ($\mathbf{D}'$) and split into m segments ($\mathbf{W_i}$) from which feature vectors ($\mathbf{X_i}$) are extracted. Given features ($\mathbf{X_i}$), a model with parameters θ scores c activity classes $\mathbf{Y_i} = \{y^1, \ldots, y^c\}$ with a confidence vector $\mathbf{p}_i$.

Our task is to teach a small child about four colors

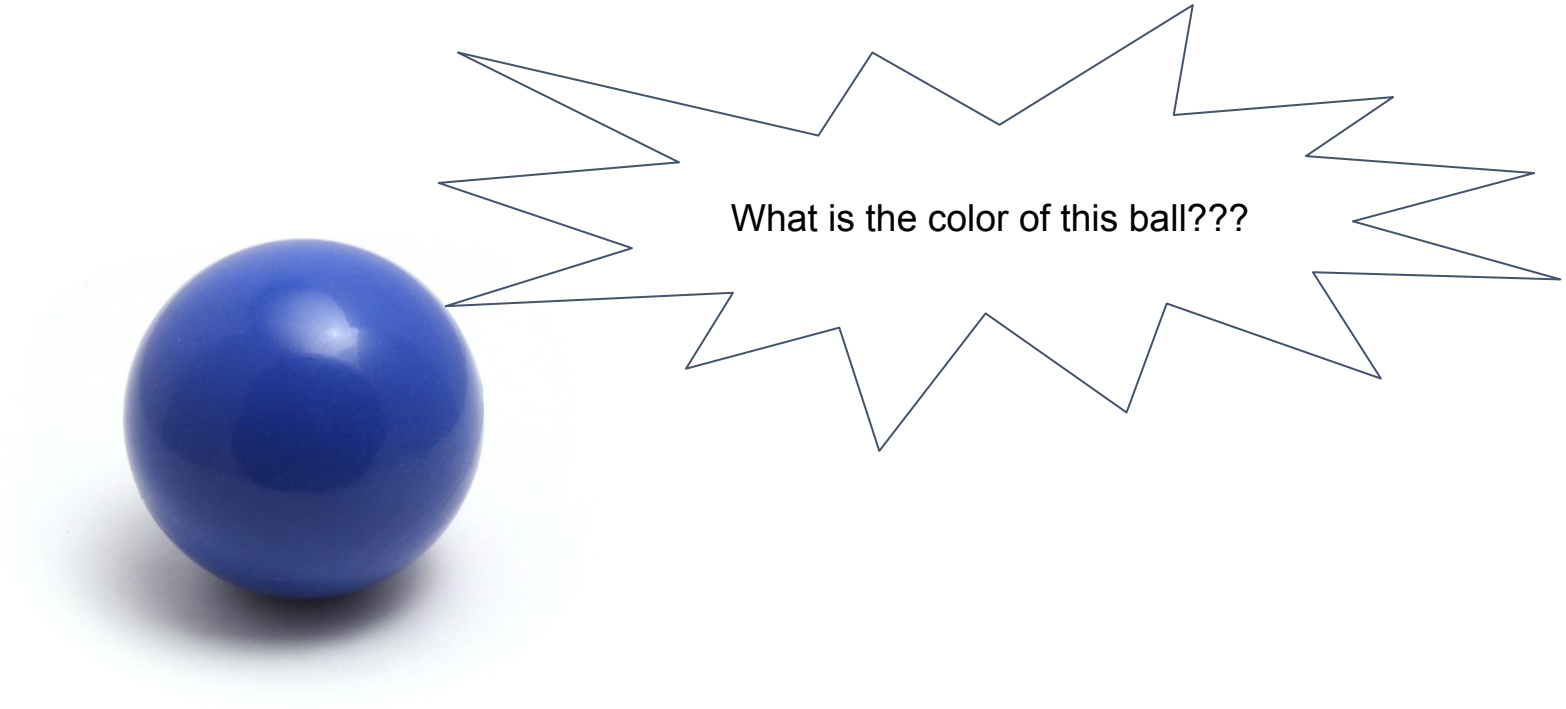
Features

How would you differentiate one ball from another?

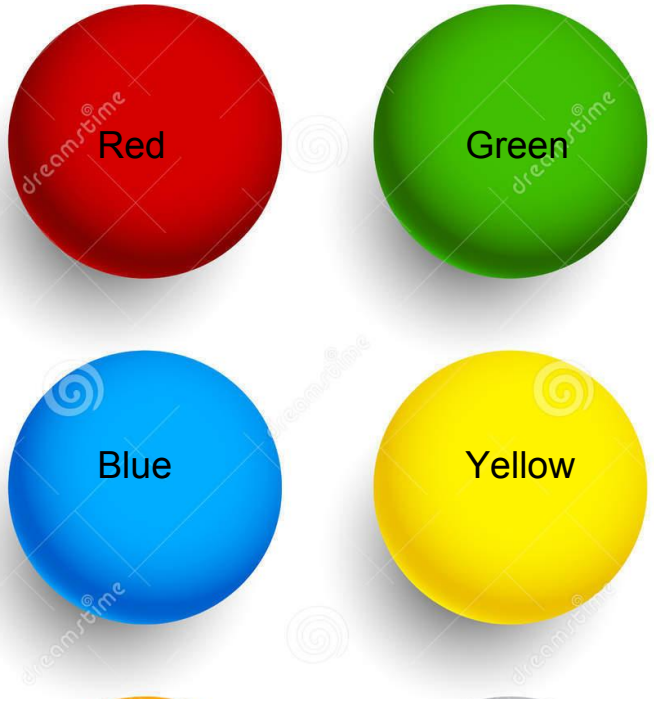
Training



Testing

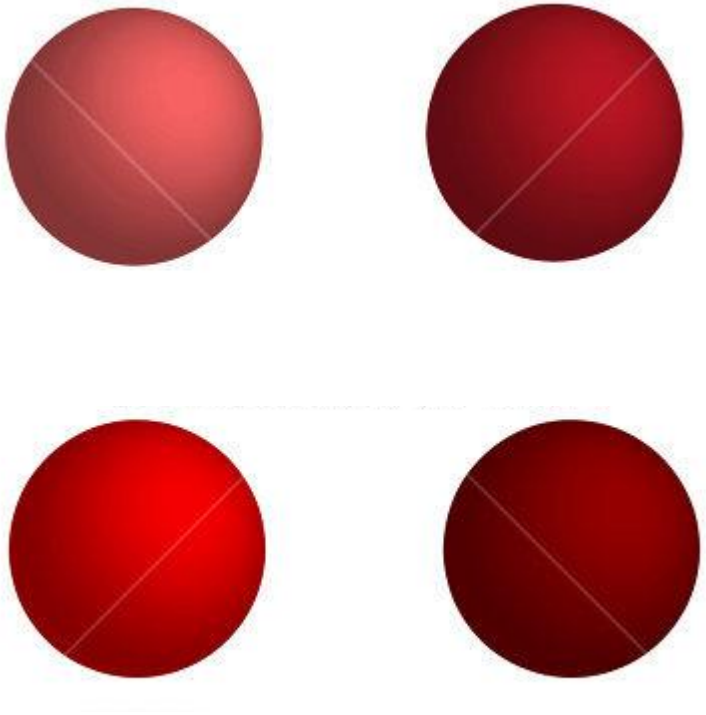


Training Data



One ball at a time!!

Testing Data



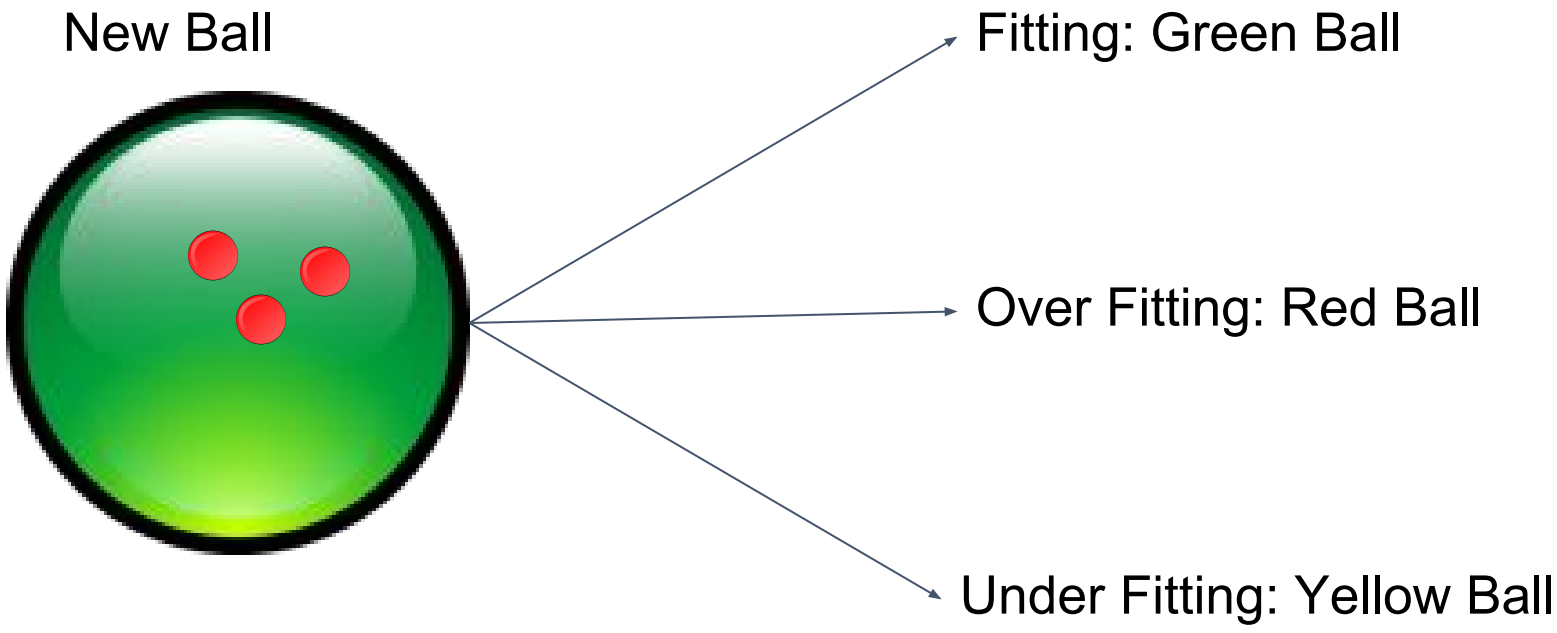
What color are they??

How to choose the best algorithm for our problem ?

Let’s say we have 20 different balls, 5 of each color(different shades)

- These 20 balls -> dataset (training + testing)
- We divide 20 balls into 5 different sets (having one of each color)
- Now randomly we select 4 sets as training dataset and one as testing (5 different times)
- What have we achieved by doing this ?

Fitting/Overfitting/Underfitting



Loss Function

New Ball

Answer: Yellow Ball

Answer: Red Ball

Metrics and errors

Metrics for Performance Evaluation…

		PREDICTED CLASS	
		Yes	No
ACTUAL CLASS	Yes	a (TP)	b (FN)
	No	c (FP)	d (TN)

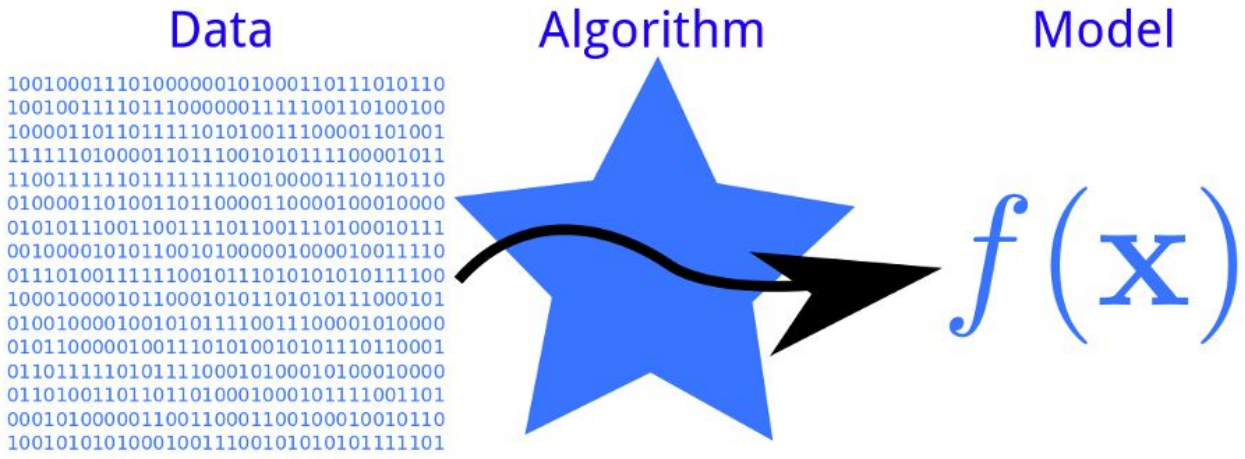
- Most widely-used metric:

$$Accuracy = \frac{a+d}{a+b+c+d} = \frac{TP+TN}{TP+TN+FP+FN}$$

Model

Take an algorithm -> Train using data -> you get a model. :)

Linear regression algorithm is a technique to fit points to a line $y = m\,x + c$. Now after fitting, you get for example, $y = 10\,x + 4$. This a model. A model is something to which when you give an input, gives an output. In ML, any 'object' created after training from an ML algorithm is a model. For example, SVM model, Random forests model, etc.

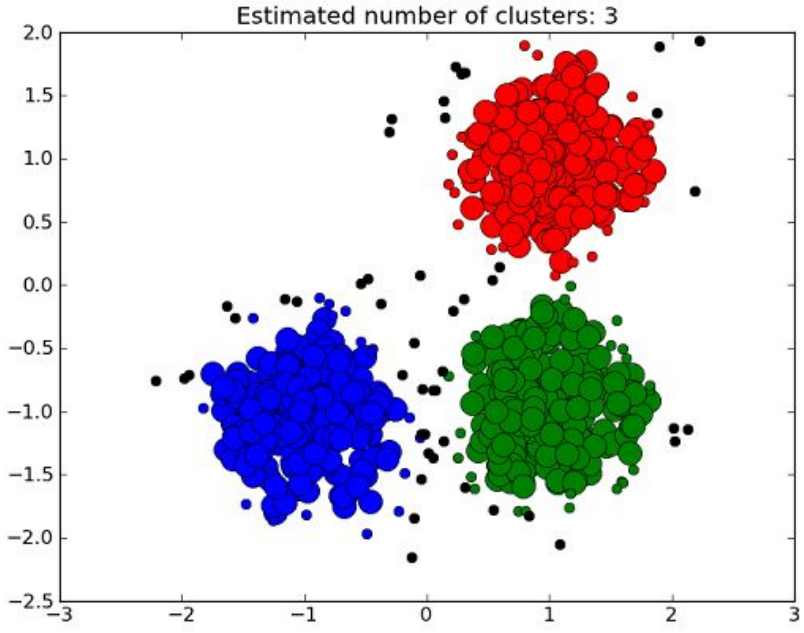


Types of Machine learning algorithms

- Unsupervised
- Supervised

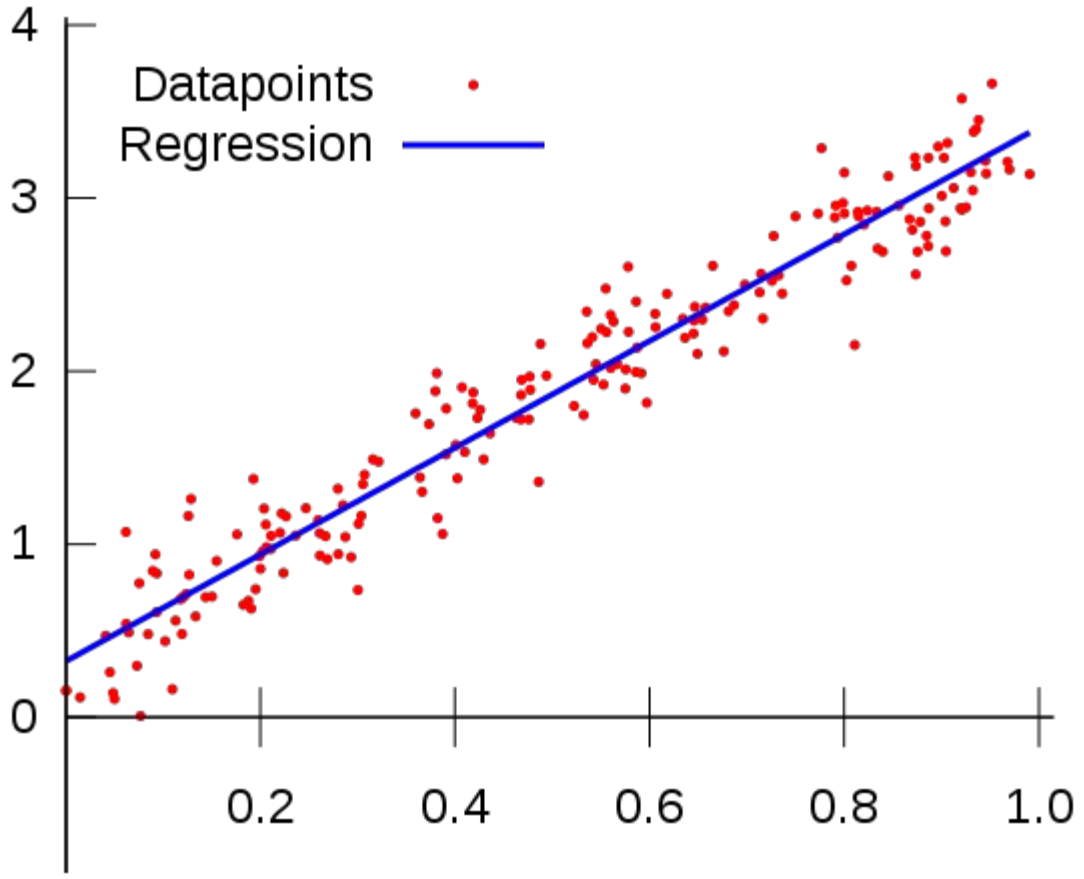
Unsupervised

- A patient reports a list of symptoms. The AI program, trained to classify symptoms into clusters, recognizes that the majority of those symptoms belong to the cluster of data that is associated with one particular disease. This results in a diagnosis of the disease for the patient.



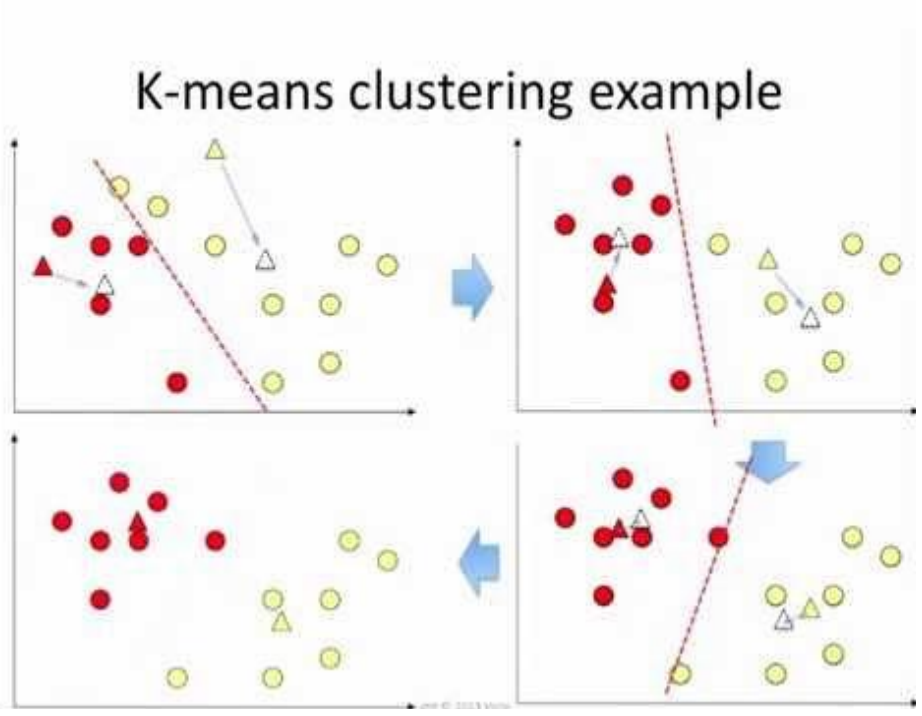
Supervised

Drug Discovery: predicting success rate based on biological factors.



Unsupervised Machine Learning

Clustering



Supervised Machine Learning

Linear Regression

Multivariate regression