# Introduction to Supervised machine learning

Introduction to machine learning

Characteristics of Supervised Machine Learning -

a. Class of machine learning algorithms that work on externally supplied instances (data) in form of predictor attributes and associated target values

b. They produce a model representing alternate hypothesis i.e. distribution of class labels in terms of predictor variables in the feature space

c. The model thus generated is used to make predictions about future instances where the predictor feature values are known but the target / class value is unknown

    a. E.g.-1 building model to predict the re-sale value of a car based on its current mileage, age, color etc.

    b. E.g.-2 Predicting the final year scores based on student performance in previous years.

Introduction to machine learning

Data Science Machine Learning Steps -
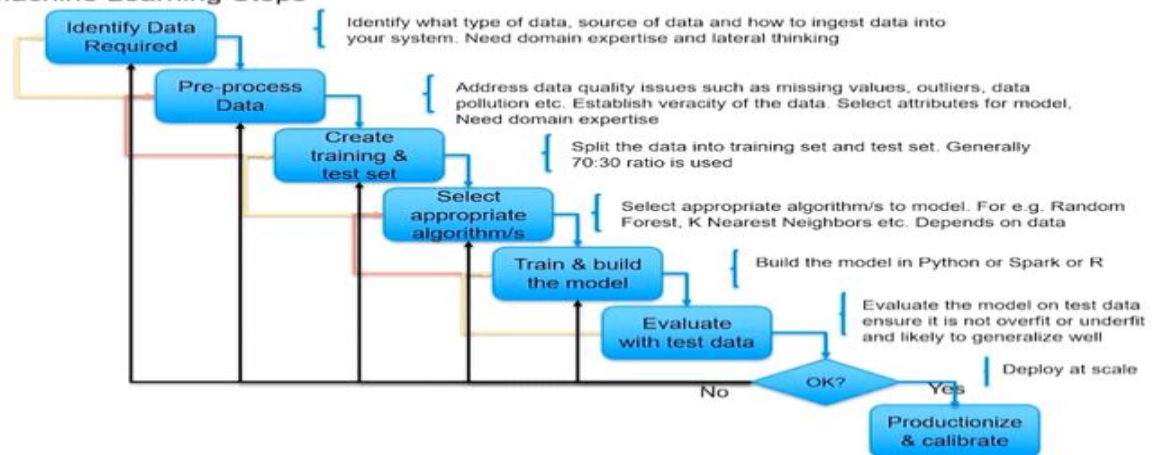
What is machine learning?

1. Is a process of enabling a computer based system to learn to do tasks based on well defined statistical and mathematical methods

2. The ability to do the tasks come from the underlying model which is the result of the learning process. Sometimes the ability comes from an mathematical algorithm

3. The model generated represents behaviour of the processes that were earlier performed before machine learning

4. The model is generated from huge volume of data, huge both in breadth and depth reflecting the real world in which the processes are performed

5. The more representative data is of the real world, the better the model would be. The challenge is how to make it a true representative

What do machine learning algorithms do?

1. Search through data to look for patterns

2. Patterns in form of trends, cycles, associations, classes etc.

3. Express these patterns as mathematical structures such as probability equations or polynomial equations

When is machine learning useful ?

1. Cannot express our knowledge about patterns as a program. For e.g. Character recognition or natural language processing

2. Do not have an algorithm to identify a pattern of interest. For e.g. In spam mail detection

3. Too complex and dynamic. For e.g. Weather forecasting

4. Too many permutations and combinations possible. For e.g. Genetic code mapping

5. No prior experience or knowledge. For e.g. Mars rover

6. Patterns hidden in humongous data. For e.g. Recommendation system

Where are machine learning based systems used (examples only)

1. Fraud detection

2. Sentiment analysis

3. Credit risk management

4. Prediction of equipment failures

5. New pricing models / strategies

6. Network intrusion detection

7. Pattern and image recognition

8. Email spam filtering

Machine Learning Pre-requisites

1. Rich set of data representing the real world

2. Knowledge and skills in
   a. Maths and statistics
   b. Programming (Python, R, Java, Go)
   c. Tools / frameworks such as Keras / TensorFlow
   d. Domain knowledge