

Machine Learning Algorithm

Author, Mayank Tripathi

A Data Science Foundation Blog

December 2019

www.datascience.foundation

Copyright 2016 - 2017 Data Science Foundation

Just a recap from my [last post](#), machine learning algorithms use the patterns contained in the training data to perform classification and future predictions. Whenever a new input is introduced to the ML model, ML applies its experience or learned patterns or data on which it trained, over the new data to make future predictions. Based on the final accuracy, one can optimize their models using various standardized approaches. In this way, Machine Learning model learns to adapt to new examples and produce better results going forward, and so on. Also it's not necessary that we always provide the data to machines. At-times machine learns by its own. Surprised! Don't be... let's see in detail.

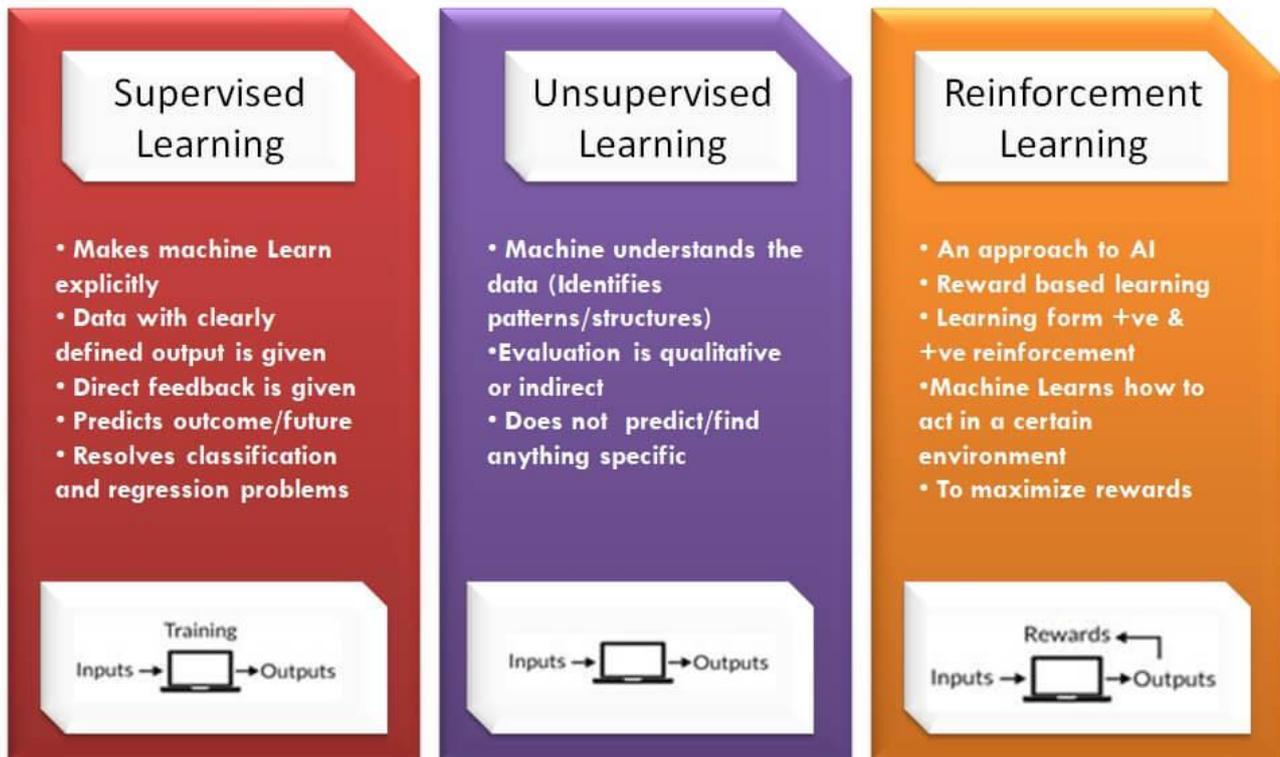
Types of Machine Learning

Machine Learning Algorithms can be classified into 3 types as

1. Supervised Learning
2. Unsupervised Learning
3. Reinforcement Learning

I don't want to add more confusion by providing too much of data rightway, as for now it may not be relevant. But just understand the basic difference between all 3. We will mainly focus on Supervised Learning. And probably once we are comfortable we can switch to other types.

Types of Machine Learning – At a Glance



Supervised learning

- In Supervised Learning, the dataset is provided, on which we train our model and it is labeled ie target is defined.
- There is a clear and distinct mapping of input / features and output / target.
- Based on the data or input(s) provided to Machine, the model is able to get trained in the instances.
- An example of supervised learning is spam / junk email filtering. Refer my previous post for details.
- Also consider the example of a Kid getting trained to recognize the fruit. This is an easier form of training out of other types.

Unsupervised Learning

- In Unsupervised Learning, there is no labeled ie target is not defined.
- Unsupervised learning is where you only have input data or features and no corresponding output or target variables.
- The algorithm identifies the patterns within the dataset and learns them.

- The algorithm groups the data into various clusters based on their density.
- These are called unsupervised learning because unlike supervised learning there is no correct answers and there is no teacher to verify it.
- Algorithms are left to their own devices to discover and present the interesting structure in the data.
- The learning process in Unsupervised Learning is solely on the basis of finding patterns in the data. After learning the patterns, the model then makes conclusions.
- Example, Movie group by the rating given by movies viewers.

Reinforcement Learning

- A system or machine which learns with the error.
- No dataset is provided, so no features or target available.
- It is all about making decisions sequentially.
- In other words we can say that the output depends on the state of the current input and the next input depends on the output of the previous input.
- Example for Reinforcement Learning, is Chess Game. Here Machine learns based on the moves of the user. Another is Robots or Self Driving Cars.

There are many Machine Learning Algorithms, but the most common and frequently used in Supervised Machine Learning are as :

1. **Regression** : This is the same as from Statistical Regression. Regression is a set of statistical processes for estimating the relationships between a dependent variable (often called the 'outcome variable') and one or more independent variables (often called 'predictors', 'covariates', or 'features'). Source : Wikipedia
The most common usage of Linear Regression where there is a linear relationship or correlation between the feature variable(s) and the target variable. Other type of Regression such as Linear Regression; Logistic Regression ARIMA Regression, each used based on the requirement.
2. **Decision Tree** : This is more popular in Predictive Modeling. In here we do create a decision tree which is able to make a decision based on user inputs.
3. **Supervised Vector Machine** : These are used to classify the data into groups or classes or categories.

After these main algorithm comes some advanced algorithm. Which can be explored later.
For better understanding of

Reference :

Wikipedia; Google; Google Images

About the Data Science Foundation

The Data Science Foundation is a professional body representing the interests of the Data Science Industry. Its membership consists of suppliers who offer a range of big data analytical and technical services and companies and individuals with an interest in the commercial advantages that can be gained from big data. The organisation aims to raise the profile of this developing industry, to educate people about the benefits of knowledge based decision making and to encourage firms to start using big data techniques.

Contact Data Science Foundation

Email: admin@datascience.foundation

Telephone: 0161 926 3641

Atlantic Business Centre

Atlantic Street

Altrincham

WA14 5NQ

web: www.datascience.foundation

Data Science Foundation

Data Science Foundation, Atlantic Business Centre, Atlantic Street, Altrincham, WA14 5NQ

Tel: 0161 926 3670 Email: admin@datascience.foundation Web: www.datascience.foundation

Registered in England and Wales 4th June 2015, Registered Number 9624670