# First understand what Machine Learning is

In simple words, Machine Learning is all about making the computers to perform intelligent tasks without explicitly coding. This is achieved by training the computer with lots of data.

# Math Skills

Under math skills, we need to know probability and statistics, linear algebra and calculus.

1. Probability and Statistics

Machine Learning is very much closely related to statistics.

You need to know:

* Fundamentals of statistics and probability theory
* Descriptive statistics
* Bayes rule
* Random variables
* Probability distributions
* Sampling
* Hypothesis testing
* Regression
* Decision analysis

1. Linear Algebra

You need to know working

* Matrices
* some basic operations on matrices such as matrix addition, subtraction, scalar and vector multiplication, inverse, transpose and vector spaces.

1. Calculus

Basics of differential and integral calculus.

# Programming Skills

A little bit of coding skills is enough. But it is preferred to have the knowledge of data structures. Algorithms and oops concepts.

Some of the popular programming language of learn for machine learning are Python, R, Jave and C++.

It is your preferred to master any one programming language.

But it is advisable to have a little understanding of other languages and what their advantages and disadvantages are over your preferred one.

# Data Engineer Skills

Ability to work with large amounts of data, data preprocessing, knowledge of SQL and NOSQL, ETL(extract transform and load) operations, Data analysis and visualization skills.

# Knowledge of Machine Learning Algorithms

|  |  |  |
| --- | --- | --- |
|  | Unsupervised | Supervised |
| Continuous | * Clustering & Dimensionally Reduction  1. SVD 2. PCA 3. K-means | * Regression  1. Linear 2. Polynomial |
| Categorical | * Association Analysis  1. Apriori 2. FP-Growth  * Hidden Markov Model | * Classification  1. KNN 2. Trees 3. Logistic Regression 4. Naïve-Bayes 5. SVM |

Should be familiar with popular Machine Learning Algorithms such as:

* Linear Regression
* Logistic Regression
* Decision Trees
* Random Forest
* Clustering: K-means
* Clustering: Hierarchical
* Reinforcement Learning
* Neural Networks

# Knowledge of Machine Learning Frameworks

Familiar with popular Machine Learning frameworks such as:

1. SCIKIT LEARN
2. TENSORFLOW
3. AZURE
4. CAFFE
5. THEANO
6. SPARK
7. TORCH