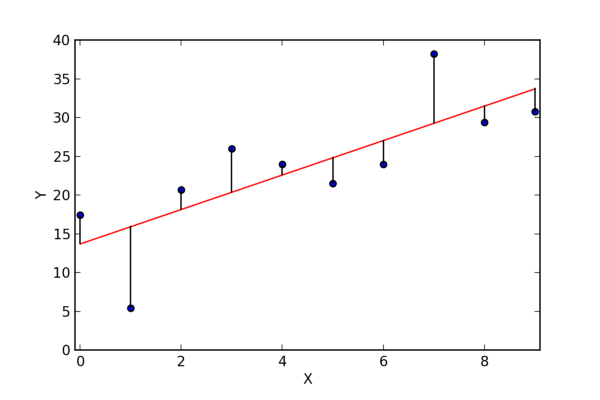
机器学习的基本算法-每个工程师必备

Machine learning as a field has been around for a long time before deep neural networks took over the scene. Here are a list of the algorithms you need to know, so you can tackle any problem that comes your way. This isn’t an exhaustive list, but your bases will be mostly covered.

Also wanted to announce that my medium blog will be transitioning from general ML focus to a deep learning focus. Most of my work nowadays involves creating novel deep learning systems, so I want to spend more time writing about it!

Regression Algorithms



Regression algorithms model relationships between variables. Originally a technique from statistics they have become an important tool in every Machine learning engineer’s tool kit.

Common Regression algorithms

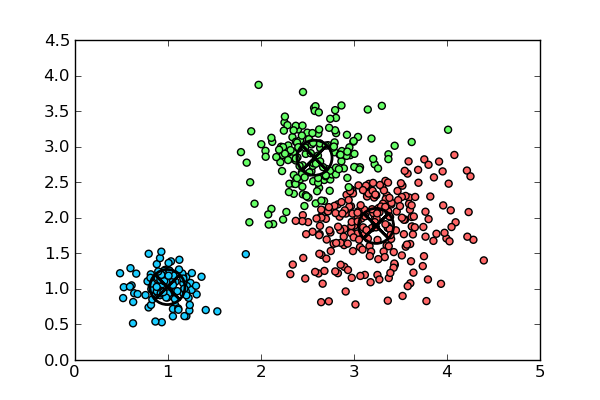
Least Squares Regression

Linear Regression

Logistic Regression

[Coursera Course](https://www.coursera.org/learn/regression-models/lecture/Kz1eV/introduction-to-regression) by Johns Hopkins on regression models

Clustering Algorithms



Clustering algorithms can divide data points in to groups with similar properties.They work by finding inherent structures in data to best organize data in to distinct groups. Things in the group are more closely related to each other then things in other groups.

There are two types of clustering algorithms. hard clustering refers to when a data point is in a group or not. soft clustering refers to when a data point can belong to many different groups to different degrees.

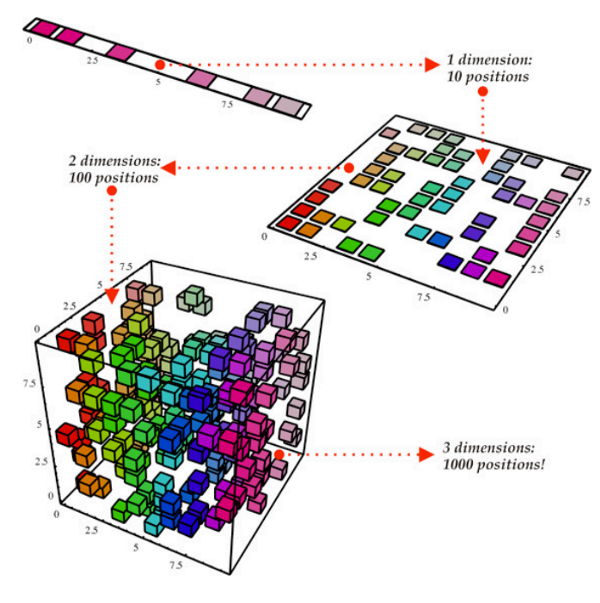
Common Clustering algorithms

K-means

Hierarchical Clustering

[Amazing introductory video on clustering](https://www.youtube.com/watch?v=ZueoXMgCd1c)

**Dimensionality reduction algorithms**



When the number of features is very large compared to the number of data points you have. Dimensional reduction algorithms help you reduce the number of features to only what is necessary for the problem at hand. They can remove redundant or useless features, helping you get better results.

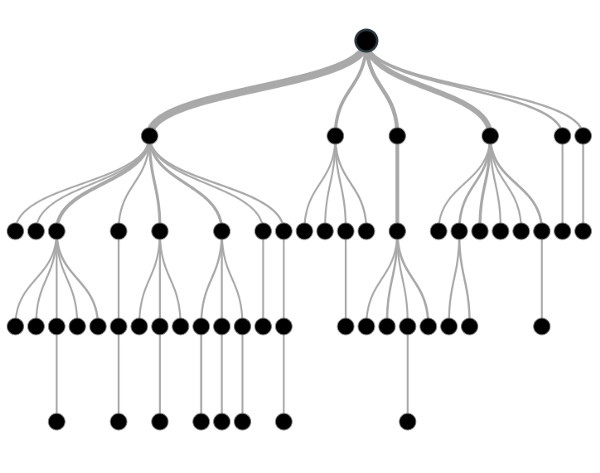
There are two ways that dimensional reduction algorithms work. The first method is through feature selection, where the algorithm picks a subset of the available features. The second way is feature extraction, which reduces the data in a high dimensional space to a lower one.

Common Dimensionality reduction algorithms

* Principle component analysis
* Low Variance Filter
* High Correlation Filter
* Random Forests
* Backward Feature Elimination / Forward Feature construction

This is not a exhaustive list, just some that I have used. If you want to read up on this some more as well as see the ROI for some of these algorithms check out [KDnuggets blog post](https://www.kdnuggets.com/2015/05/7-methods-data-dimensionality-reduction.html) on it.

**Decision tree algorithms**



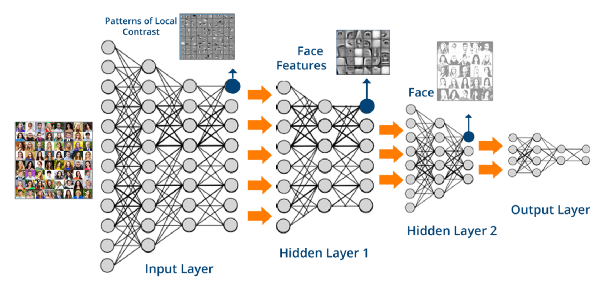
decision trees create models of decisions made on values from your data. A fork is made in the tree structure until there is a prediction for every data point. Their results are easy to understand unlike other algorithms (Deep Learning) and they are easy to use on many different data types.

Common decision tree algorithms:

* Classification and Regression Tree
* C4.5 and C5.0
* Random Forests
* Chi-squared automatic interaction detector

[Analytics Vidhya has a great article](https://www.analyticsvidhya.com/blog/2016/04/complete-tutorial-tree-based-modeling-scratch-in-python/) that goes in depth on decision trees. Listing out the different algorithms and their advantages and disadvantages

**Deep Learning**



The hype behind machine learning and “AI” is caused by deep learning. They are modern versions of artificial neural networks that exploit cheap computation to train ever larger neural networks. They are powerful universal function approximates that have proven their ability in solving some of the hardest problems. See [Alpha Go](https://en.wikipedia.org/wiki/AlphaGo).

Common Deep learning al

* Stacked Auto-encoders
* Convolution Neural networks
* Recurrent neural networks
* Capsule Networks ([more information here](https://hackernoon.com/what-is-a-capsnet-or-capsule-network-2bfbe48769cc))

[Check out this book snippet.](https://www.safaribooksonline.com/library/view/deep-learning/9781491924570/ch04.html) It goes over the major architectures for deep learning.

**Take away**

If you serious about machine learning you have to understand the tools that are available to you. Having a good understanding of these tools will give you a leg up on any problems you come across.

关键词：Machine Learning；Programming；Tutorial