

Search Data Science Central [Search](#)

- [Ramiro Arce](#)
- [Sign Out](#)



# Data Science Central™

THE ONLINE RESOURCE FOR BIG DATA PRACTITIONERS

• [HOME](#) [DATAVIZ](#) [HADOOP](#) [BIG DATA](#) [ANALYTICS](#) [WEBINARS](#) [DEEP LEARNING](#) [AI](#) [JOBS](#) [MEMBERSHIP](#) [SEARCH](#) [CLASSIFIEDS](#) [CONTACT](#)

[Subscribe to DSC Newsletter](#)

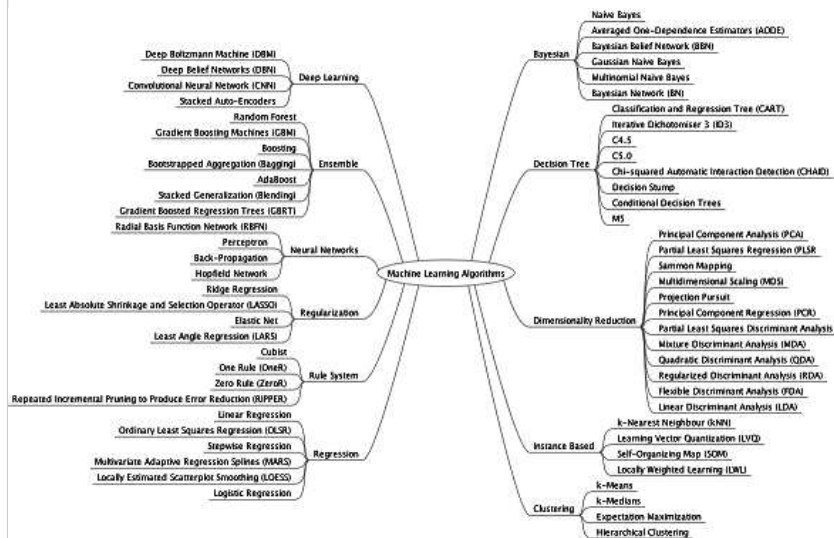
- All Blog Posts
- My Blog
- Edit Blog Posts
- Add



## A Tour of Machine Learning Algorithms

- Posted by L.V. on October 22, 2015 at 9:30am
- [Send Message](#) [View Blog](#)

Originally published by Jason Brownlee in 2013, it still is a goldmine for all machine learning professionals. The algorithms are broken down in several categories. Here we provide a high-level summary, a much longer and detailed version can be found [here](#). You can even download an algorithm map from the original article. Below is a much smaller version.

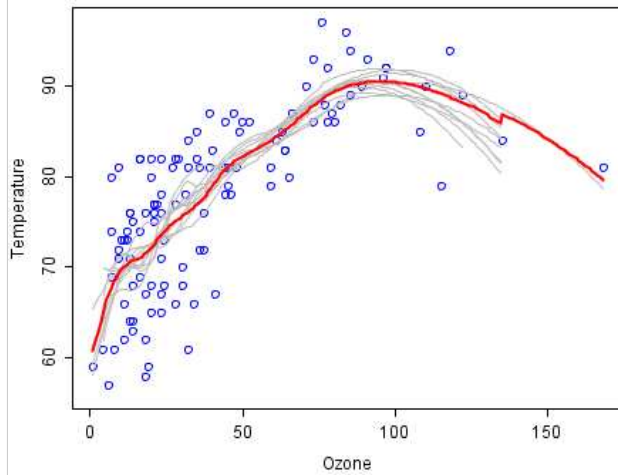


It would be interesting to list, for each algorithm,

- examples of real world applications,
- in which contexts it performs well,
- if it can be used as a black box,
- ease of use and interpretation,
- how it handles missing data,
- enterprise version available or not,
- integration with existing analytics platforms or real-time systems,
- constraints on data (e.g. Naive Bayes performs poorly on correlated variables),
- maintenance/scalability issues,
- distributed implementation,
- speed or computational complexity,
- can easily be blended with other algorithms

and generally speaking, compare these algorithms. I would add HDT, Jackknife regression, density estimation, attribution modeling (to optimize marketing mix), linkage (in fraud detection), indexation (to create taxonomies or for clustering large data sets consisting of text), bucketisation, and time series algorithms.

For more on machine learning (ML), [click here](#).



*Ensemble methods to fit data: see original paper*

#### 1. Regression Algorithms

- Ordinary Least Squares Regression (OLSR)
- Linear Regression
- Logistic Regression
- Stepwise Regression
- Multivariate Adaptive Regression Splines (MARS)
- Locally Estimated Scatterplot Smoothing (LOESS)

#### 2. Instance-based Algorithms

- k-Nearest Neighbour (kNN)
- Learning Vector Quantization (LVQ)
- Self-Organizing Map (SOM)
- Locally Weighted Learning (LWL)

#### 3. Regularization Algorithms

- Ridge Regression
- Least Absolute Shrinkage and Selection Operator (LASSO)
- Elastic Net
- Least-Angle Regression (LARS)

#### 4. Decision Tree Algorithms

- Classification and Regression Tree (CART)
- Iterative Dichotomiser 3 (ID3)
- C4.5 and C5.0 (different versions of a powerful approach)
- Chi-squared Automatic Interaction Detection (CHAID)
- Decision Stump
- M5
- Conditional Decision Trees

#### 5. Bayesian Algorithms

- Naive Bayes
- Gaussian Naive Bayes
- Multinomial Naive Bayes
- Averaged One-Dependence Estimators (AODE)
- Bayesian Belief Network (BBN)
- Bayesian Network (BN)

#### 6. Clustering Algorithms

- k-Means
- k-Medians
- Expectation Maximisation (EM)
- Hierarchical Clustering

#### 7. Association Rule Learning Algorithms

- Apriori algorithm
- Eclat algorithm

#### 8. Artificial Neural Network Algorithms

- Perceptron
- Back-Propagation
- Hopfield Network
- Radial Basis Function Network (RBFN)

#### 9. Deep Learning Algorithms

- Deep Boltzmann Machine (DBM)
- Deep Belief Networks (DBN)
- Convolutional Neural Network (CNN)
- Stacked Auto-Encoders

#### 10. Dimensionality Reduction Algorithms

- Principal Component Analysis (PCA)
- Principal Component Regression (PCR)
- Partial Least Squares Regression (PLSR)
- Sammon Mapping
- Multidimensional Scaling (MDS)
- Projection Pursuit
- Linear Discriminant Analysis (LDA)
- Mixture Discriminant Analysis (MDA)
- Quadratic Discriminant Analysis (QDA)
- Flexible Discriminant Analysis (FDA)

11. Ensemble Algorithms

- Boosting
- Bootstrapped Aggregation (Bagging)
- AdaBoost
- Stacked Generalization (blending)
- Gradient Boosting Machines (GBM)
- Gradient Boosted Regression Trees (GBRT)
- Random Forest

12. Other Algorithms

- Computational intelligence (evolutionary algorithms, etc.)
- Computer Vision (CV)
- Natural Language Processing (NLP)
- Recommender Systems
- Reinforcement Learning
- Graphical Models

DSC Resources

- Career: Training | Books | Cheat Sheet | Apprenticeship | Certification | Salary Surveys | Jobs
- Knowledge: Research | Competitions | Webinars | Our Book | Members Only | Search DSC
- Buzz: Business News | Announcements | Events | RSS Feeds
- Misc: Top Links | Code Snippets | External Resources | Best Blogs | Subscribe | For Bloggers

Additional Reading

- 50 Articles about Hadoop and Related Topics
- 10 Modern Statistical Concepts Discovered by Data Scientists
- Top data science keywords on DSC
- 4 easy steps to becoming a data scientist
- 13 New Trends in Big Data and Data Science
- 22 tips for better data science
- Data Science Compared to 16 Analytic Disciplines
- How to detect spurious correlations, and how to find the real ones
- 17 short tutorials all data scientists should read (and practice)
- 10 types of data scientists
- 66 job interview questions for data scientists
- High versus low-level data science

Follow us on Twitter: @DataScienceCtrl | @AnalyticBridge

Views: 94499

Like  
36 members like this

Share Tweet  Facebook

Like 4

- < Previous Post
- Next Post >

Comment



Visual Mode

HTML Editor



10pt

**B**

*I*

~~S~~

U

*I*<sub>x</sub>

Follow – Email me when people comment

Add Comment



Comment by Ben Dutta on December 29, 2015 at 2:58pm