Algorithm	Learning type	Class	Restriction bias	Preference bias
K-Nearest Neighbors	Supervised	Instance based	Generally speaking, KNN is good for measuring distance-based approximations; it suffers from the curse of dimensionality	Prefers problems that are distance based
Naive Bayes	Supervised	Probabilistic	Works on problems where the inputs are independent from each other	Prefers problems where the probability will always be greater than zero for each class
Decision Trees/ Random Forests	Supervised	Tree	Becomes less useful on problems with low covariance	Prefers problems with categorical data
Support Vector Machines	Supervised	Decision boundary	Works where there is a definite distinction between two classifications	Prefers binary classification problems
Neural Networks	Supervised	Nonlinear functional approximation	Little restriction bias	Prefers binary inputs
Hidden Markov Models	Supervised/ Unsupervised	Markovian	Generally works well for system information where the Markov assumption holds	Prefers time-series data and memoryless information
Clustering	Unsupervised	Clustering	No restriction	Prefers data that is in groupings given some form of distance (Euclidean, Manhattan, or others)
Feature Selection	Unsupervised	Matrix factorization	No restrictions	Depending on algorithm can prefer data with high mutual information
Feature Transformation	Unsupervised	Matrix factorization	Must be a nondegenerate matrix	Will work much better on matricies that don't have inversion issues
Bagging	Meta-heuristic	Meta-heuristic	Will work on just about anything	Prefers data that isn't highly variable

Symbol	How do you say it?	What does it do?
$\sum_{i=0}^{n} x_i$	The sum of all *x*s from x_0 to x_n	This is the same thing as $x_0 + x_1 + \cdots + x_n$.
$\|x\|$	The absolute value of *x*	This takes any value of *x* and makes it positive. So $\|-x\| = \|x\|$.
$\sqrt{4}$	The square root of 4	This is the opposite of 2^2.
$z_k = < 0.5, 0.5 >$	Vector z_k equals 0.5 and 0.5	This is a point on the xy plane and is denoted as a vector, which is a group of numerical points.
$log_2(2)$	Log 2	This solves for *i* in $2^i = 2$.
$P(A)$	Probability of A	In many cases, this is the count of A divided by the total occurrences.
$P(A\|B)$	Probability of A given B	This is the probability of A and B divided by the probability of B.
$\{1,2,3\} \cap \{1\}$	The intersection of set one and two	This turns into a set {1}.
$\{1,2,3\} \cup \{4,1\}$	The union of set one and two	This equates to {1,2,3,4}.
$det(C)$	The determinant of the matrix C	This will help determine whether a matrix is invertible or not.
$a \propto b$	a is proportional to b	This means that $m \cdot a = b$.
min $f(x)$	Minimize $f(x)$	This is an objective function to minimize the function $f(x)$.
X^T	Transpose of the matrix *X*	Take all elements of the matrix and switch the row with the column.