

*ComparisonofDataMiningtechniques–largedatasets–Guidelines(...andonlyguidelines)**H:high,M:medium,L:low.*

	Multiple Linear Regression	Logistic Regression	Discriminant Analysis	Naïve Bayes	Neural Nets	Trees	k-Nearest Neighbors
Accuracy	M	M	M	HM	H	M	HM
Intepretability	H	H	M	H	L	H	L
Speed-Training	H	H	H	HM	L	HM	H
Speed-Deployment	H	H	H	H	H	HM	L
Effort in choice and transformation of indep. Vars.	HM	HM	HM	HM	L	L	ML
Effort to tune performance parameters	L	L	L	ML	H	ML	ML
Robustness to Outliers in indep vars	ML	ML	ML	ML	HM	H	HM
Robustness to irrelevant variables	H	H	HM	H	L	ML	L
Ease of handling of missing values	M	M	M	H	M	H	ML
Natural handling both categorical and continuous variables	H	H	ML	M	H	H	L