Which Machine Learning Algorithm Is Right for You?

Algorithm	Dataset	Training Speed	Interpretability	Tuning	Comments
	What is the ideal dataset size for each algorithm?	*How quickly will the algorithm train without acceleration hardware?*	*How hard is it to see how the algorithm arrived at a decision?*	*How much tuning does the algorithm allow?*	
Linear models	Small	Very fast	Easy	Minimal	Widely used basic algorithm Linear SVM handles high-dimensional data well
Decision trees	Small	Very fast	Easy	Some	Good generalist algorithm, check for overfitting
(Nonlinear) Support vector machine	Medium sized	Moderately slow	Difficult	Some	Good accuracy
Nearest neighbor	Medium sized	Moderately fast	Moderately easy	Minimal	Lower accuracy, but easy to use and interpret
Naïve Bayes	Medium sized	Very fast	Moderately easy	Some	Widely used for text analytics (e.g., spam filtering); kernel Bayes will run slower
Ensembles	Large	Moderately fast	Difficult	Some	Higher accuracy with a tradeoff of lower interpretability
Neural network (shallow)	Medium sized	Moderately fast	Moderately easy	Some	Still used for signal classification, compression, and forecasting
Deep nets	Large	Very slow	Difficult	A lot	A standard algorithm for image, video, signals, and text