

Comparison of five classification algorithms

Comparison of algorithms	Type of input acceptable	Type of output that can be predicted.	Can allocate an example to more than one class	Output easily understood	Handles errors (missing data / outliers)	Scalable	Fast training time?	Predict complex relationships?
Decision Tree	Categorical only, but some implementations can take numeric input.	Categorical only.	No	Yes	Sensitivity to noise configurable	No	Computationally inexpensive. Irrelevant attributes slow down training time.	No
Neural Network	Most efficient with normalised numeric input, but can handle categorical as well.	Categorical or numeric.	No	No	Can slow down training time	Somewhat	Can be slow depending on complexity of the pattern and initial network configuration.	Yes
Nearest Neighbour	All (provided distance measure used support categorical data)	Categorical	No	No	Sensitivity to noise is dependent on value of K.	No	Yes , but subsequent classification can be slow.	Good for datasets with a lot of variability. Accuracy depends on appropriate value for K
Support Vector Machine	Numeric only	Binary class label	No	No	Sensitivity to noise configurable	Yes	Yes , training time not relative to size of dataset.	Need linear decision boundary
Naïve Bayes	All, but preferably categorical, attributes.	Categorical.	Yes	OK	Yes	No	Can be slow	No