eCornell

CHEAT SHEET

Bayes Optimal Classifier

Algorithm Name	Bayes Optimal Classifier
Description	Given $P(y\|\mathbf{x})$, the Bayes Optimal Classifier selects the most probable label for each test point.
Applicability	Useful for benchmarking other classifiers if the true distribution of the data is known.
Assumptions	The true distribution of the data is known.
Underlying Mathematical Principles	Risk minimization is used to derive the Bayes Classifier, although it's relatively intuitive to simply select the most probable label for each $\mathbf{x}$. You make a prediction by finding the arguments of the maxima (arg max) of the $P(y\|\mathbf{x})$. The predicted value optimizes a function $h(\mathbf{x})$. $y^* = h_{opt}(\mathbf{x}) = \arg\max_y \mathrm{P}(y\|\mathbf{x})$ If a test data point does not have the most likely label, the prediction will be incorrect. You can computer the error rate exactly: $\epsilon_{BayesOpt} = 1 - \mathrm{P}(h_{\mathrm{opt}}(\mathbf{x})\|\mathbf{x}) = 1 - \mathrm{P}(y^*\|\mathbf{x})$
Additional Details	Use `numpy.argmax (p_labels)` to find the label that has the greatest probability for a given $\mathbf{x}$.