
Ilya Narsky†

California Institute of Technology
(Dated: July 19, 2005)

Abstract
Modern analysis of high energy physics (HEP) data needs advanced statistical tools to separate signal from background. A C++ package has been implemented to provide such tools for the HEP community. The package includes linear and quadratic discriminant analysis, decision trees, bump hunting (PRIM), boosting (AdaBoost), bagging and random forest algorithms, and interfaces to the standard backpropagation neural net and radial basis function neural net implemented in the Stuttgart Neural Network Simulator. Supplemental tools such as bootstrap, estimation of data moments, and a test of zero correlation between two variables with a joint elliptical distribution are also provided. The package offers a convenient set of tools for imposing requirements on input data and displaying output. Integrated in the BaBar computing environment, the package maintains a minimal set of external dependencies and therefore can be easily adapted to any other environment. It has been tested on many idealistic and realistic examples.

PACS numbers: 02.50.Tt, 02.50.Sk, 02.60.Pn.

* Work partially supported by Department of Energy under Grant DE-FG03-92-ER40701.
†Electronic address: narsky@hep.caltech.edu