

將軍

Sören Sonnenburg[†], Gunnar Rätsch[‡], Fabio De Bona[‡]



Fraunhofer Institut
Rechnerarchitektur
und Softwaretechnik



MAX-PLANCK-GESELLSCHAFT

- [†] Fraunhofer FIRST.IDA, Kekuléstr. 7, 12489 Berlin, Germany
- [‡] Friedrich Miescher Laboratory of the Max Planck Society,
Spemannstr. 37-39, 72076 Tübingen, Germany

Machine Learning Toolbox SHOGUN features algorithms:

- to train hidden markov models
- to learn regression and 2-class classification problems
- toolbox's focus is on kernel methods esp. SVMs
- also implements a number of linear methods like LDA, LPM, (Kernel) Perceptrons

FEATURES I

- SHOGUN interfaces to MatlabTM, Octave and Python-numarray and R
- provides generic SVM object interfacing to *seven* different SVM implementations, among them the state of the art LibSVM and SVM^{light}
- SVMs can be combined with a variety of common kernels (efficient impl. for: Linear, Polynomial, Gaussian and Sigmoid Kernel)
- recent string kernels as e.g. the Spectrum or Weighted Degree Kernel (with shifts) (`linadd` optimized)
- or working with custom pre-computed kernels.
- can use “combined kernel”, can be constructed by a weighted linear combination of a number of sub-kernels (not necessarily working on the same domain)

FEATURES II

- kernel weighting can be learned using Multiple Kernel Learning.
- input feature-objects can be dense, sparse or strings and of type int/short/double/char; can be converted into different feature types.
- chains of “preprocessors” (e.g. subtracting the mean) can be attached to each feature object (on-the-fly pre-processing)
- can make use of multi-processor machines (training on 10 Mio splice sites was possible!)

DEMO:

- Matlab
- Python
- R
- Octave

HELP

We need help:

- Documentation
- Examples
- Testing
- Test Suite

Source Code is freely available for academic non commercial use