

Weka[7] LibSVM

作者: Koala++/屈伟

首先要提的是 LibSVM 是一个库, Lib 很明显是 library 的缩写, 有些人不知道怎么会认为它是一种算法。它是由中国台湾的 Chih-Chung Chang 和 [Chih-Jen Lin](#) 等人开发的, 他们用多种语言实现写了 LibSVM。

我把这一篇放到 Weka 开发里讲, 主要讲它怎么和 Weka 结合, Weka 中并不是没有 SVM 算法, Weka 中有 SMO 算法的实现。

[Weka](#) and [LibSVM](#) are two efficient software tools for building SVM classifiers. Each one of these two tools has its points of strength and weakness. Weka has a GUI and produces many useful statistics (e.g. confusion matrix, precision, recall, F-measure, and ROC scores). LibSVM runs much faster than Weka SMO and supports several SVM methods (e.g. One-class SVM, nu-SVM, and R-SVM). Weka LibSVM (WLSVM) combines the merits of the two tools. WLSVM can be viewed as an implementation of the LibSVM running under Weka environment.

这一段是我从拷贝来的, 请注意里面有一句话, LibSVM 运行的比 Weka 里的 SMO 快的多, 如果你敢用 SMO 算法去训练大数据集, 你就明白天荒地老的真实含意了。其它和 Weka 结合的最主要的原因, 我认为是, 我们开始的时候往往都是用别的算法去试着做实验的(或者本来就需要多种基分类器), 比如 Naïve Bayes(LibSVM 是比 SMO 快的多, 可是比起来 Naïve Bayes, 它还是蜗牛), 到最后又想试试 LibSVM, 这时就需要 LibSVM 库。

还有一点也是容易误解的就是 Weka 里的高版本, 里面是有 LibSVM 这个分类器, 但你如果直接想运行是会出错的, 提示你没有设置路径。原因是: WLSVM can be viewed as an implementation of the LibSVM running under Weka environment.

先把 LibSVM 下载下来, 上次竟然有人问我在哪下载, 这种问题, 我真是不想回答, 不就是 google 一下吗? 网址也放上: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>, 下载 WLSVM, 解压后 Lib 文件夹下有一个 LibSVM.jar 的包, 用和导入 Weka.jar 包相同的方式导入就好了, 然后使用 LibSVM 和使用以前任何一种分类器的方式都是一样的。

```
LibSVM classifier = new LibSVM();
classifier.buildClassifier( instances );

Evaluation evaluation = new Evaluation( ins );
evaluation.crossValidateModel(classifier, ins, 10, random );

System.out.println( "正确率为: " + ( 1 - evaluation.errorRate() ) );
```