# О реализациях алгоритмов BMRM и LS-BMRM.

У ряда алгоритмов обучения по прецедентам задача обучения может быть поставлена в виде задачи минимизации регуляризованного эмпирического риска (regularized risk minimization, [], с. 7). Для ее решения разработаны алгоритм BMRM (Bundle Method For Regularized Risk Minimization) и его модификация LS-BMRM ([], []).

## Существующие реализации

Доступны следующие пакеты с открытым кодом:

* BMRM (<http://users.cecs.anu.edu.au/~chteo/BMRM.html>)
* LIBOCAS (<http://cmp.felk.cvut.cz/~xfrancv/ocas/html/index.html>) – реализации линейного SVM-классификатора (OCAS-бинарный, OCAM-мультиклассовый классификатор), при обучении используется алгоритм LS-BMRM. Также код может быть найден на странице SHOGUN machine learning toolbox (<http://www.shogun-toolbox.org/>).

## Эксперименты

* В работе [2] сравниваются реализации следующих алгоритмов оптимизации для линейных SVM-классификаторов: BMRM, LS-BMRM (оба из пакета BMRM) и BT (Bundle Method типа «trust region»). Критерий сравнения - число итераций, за которое алгоритм находит приближенное решение с заранее определенной точностью. Как отмечают авторы, на всех наборах данных LS-BMRM показал лучший результат. В основном BMRM выигрывает у BT на данных с большим количеством признаков и проигрывает в остальных случаях.
* В [2] также приведен эксперимент, в котором сравнивается время работы бинарных линейных SVM-классификаторов (отличаются алгоритмами оптимизации). Использовались реализации:
  + Пакет BMRM – алгоритм BMRM.
  + OCAS – алгоритм LS-BMRM.
  + LIBLINEAR v1.33 – dual coordinate descent method (Hsieh, 2008).
  + SVMperf v2.5 (Joachims, 2006).

По скорости работы реализации располагаются следующим образом:

1. LIBLINEAR
2. OCAS
3. BMRM
4. SVMperf

Отмечается значительное превосходство LIBLINEAR и OCAS над BMRM и SVMperf.

* В работе [3] OCAS сравнивается с другими бинарными линейными SVM классификаторами:
  + SGD, Pegasos - *approximative solvers,* используют эвристические критерии останова при решении оптимизационной задачи обучения.
  + SVMlight , SVMperf , BMRM, GPDT - *accurate solvers,* обучение продолжается, пока не будет достигнута требуемая точность решения.

Также в [3] сравниваются две реализации мультиклассового линейного SVM-классификатора: OCAM (алгоритм LS-BMRM) и CPAM (алгоритм BMRM). По всем рассматриваемым параметрам OCAM выглядит лучше, чем CPAM.

* В [] сравниваются SVMLight, SVMperf, Pegasos, DCDM (Dual Coordinate Descent Method), BMRM на примере задачи распознавания языка. Среди прочего авторы отмечают следующее:
  + Для всех алгоритмов время обучения линейно зависит от объема обучающей выборки.
  + Все алгоритмы достигают приблизительно одинаковой обобщающей способности, имея в распоряжении достаточно времени. Алгоритмы отличаются характером сходимости и зависимостью от объема обучающей выборки.
  + BMRM – не самый быстрый алгоритм, но хорошо подходит для распределённых вычислений.

## Литература

1. V Franc, S Sonnenburg, T Werner. Cutting Plane Methods in Machine Learning. Optimization for Machine Learning, MIT Press, 2011.
2. C. Teo, S. Vishwanathan, A. Smola, and V. Quoc. Bundle methods for regularized risk minimization. *Journal of Machine Learning Research*, 11:311–365, 2010.
3. V. Franc and S. Sonnenburg. Optimized cutting plane algorithm for support vector machines. In A. McCallum and S. Roweis, editors, *Proceedings of the International Conference on Machine* *Learning*, pages 320–327. Omnipress, 2008.
4. Cumani S; Castaldo F; Laface P.; Colibro D; Vair C (2010) Comparison of Large-scale SVM Training Algorithms for Language Recognition. In: Odysssey 2010,The Speaker and Language Recognition Workshop, Brno, 28/06 - 1/07 20010.