Release Note of LIBLINEAR 2.42

Chih-Jen Lin

Department of Computer Science
National Taiwan University
cjlin@csie.ntu.edu.tw

1 Introduction

In training logistic regression and L2-loss linear SVM, LIBLINEAR provides two types of solvers

- A coordinate descent (CD) method to solve the dual problem (the default solver).
- A truncated Newton method to solve the original primal problem.

They are respectively first-order and second-order methods, and are suitable under different circumstances. In Table 1 we borrow a table in the appendix of Fan et al. (2008) to describe their properties.

From Table 1 and feedback of users, the default solver (dual CD method) may be slow in some situations (e.g., data not scaled). In the past, if slow convergence occurs, LIBLINEAR issues a warning message suggesting users to use the primal Newton method. In this release, we make such a switch automatic to ensure that a reasonably good approximate solution of the optimization problem is directly returned to the user.

2 Implementation Details and Experimental Results

We begin with presenting Table 2 to confirm the slow convergence of the dual CD method. We compare

- primal Newton, and
- dual CD

by using the regularization parameter $C = 100C_{\text{best}}$, where C_{best} is the value that leads to the best cross validation (CV) accuracy. We obtain C_{best} by using the parameter selection tool in LIBLINEAR (the `-C` option) with a small stopping tolerance 0.0001. Note that this small tolerance is only used to accurately get C_{best}. In all other experiments the default condition with a larger tolerance is applied. We run five-fold CV and present the following information.

- Number of iterations in each training process, where four of the five folds are used as the training data.
- CV accuracy.
- Total elapsed CV running time in seconds. We do not exclude the time to read data and predict each fold.

Experiments in this note were conducted on a machine with Intel Xeon E5-2620 2.00GHz CPU. Other jobs may be run on the same computer, so the running time may not be accurate. However, for the same data set, as primal and dual solvers are consecutively called and the machine load may not change much in a time period, the comparison should correctly indicate which one is faster.

From Table 2, for some problems dual CD does not satisfy its default stopping condition after 1,000 cycles of going through all variables. The returned model may give inferior performances. For example, the CV accuracy by the dual CD for problem german.numer is 5% lower than that by the primal Newton.

In this release, we lower the maximal number of CD iterations from 1,000 to a smaller number, and impose the following rule to switch to the primal Newton method.

Table 1: A table from Fan et al. (2008) to describe properties of LIBLINEAR solvers for logistic regression and L2-loss linear SVM.

	Dual-based solvers (coordinate descent methods)	Primal-based solvers (Newton-type methods)
Property	Extremely fast in some situations, but may be very slow in some others	Moderately fast in most situations
When to use it	1. Large sparse data (e.g., documents) under suitable scaling and C is (regularization parameter) not too large 2. Data with # instances $\ll$ # features	Others

If CD iterations $\geq$ limit of CD iterations
switch to primal Newton

An issue is about the initial $\boldsymbol{w}$ of the newly called primal Newton method. Two options are

1. the $\boldsymbol{w}$ returned from the unfinished dual CD, and
2. the $\boldsymbol{w} = \mathbf{0}$ used as if the primal Newton is freshly considered.

We decide to use the first option (i.e., $\boldsymbol{w}$ returned from dual CD) for the following reasons.

- Reassigning $\boldsymbol{w}$ to $\mathbf{0}$ needs additional lines of code.
- The returned $\boldsymbol{w}$ from dual CD may be a better initial point than $\mathbf{0}$ because some CD steps have been spent for training.

Note that the stopping condition of the Newton method is not affected because of the choice of the initial point. In the primal Newton method, the condition always uses $\mathbf{0}$ on the right-hand side:

$$\|\nabla f(\boldsymbol{w})\| \leq \epsilon \frac{\min\{\#\text{pos}, \#\text{neg}\}}{l} \|\nabla f(\mathbf{0})\|, \tag{1}$$

where $f(\cdot)$ is the primal problem to be minimized, #pos and #neg are the numbers of positive- and negative-labeled instances respectively, and l is the total number of training data. However, the stopping tolerance ϵ should be adjusted because it is the one used by dual CD. For example, for logistic regression, the default ϵ for dual CD is 0.1, but the default ϵ for primal Newton is 0.01. Therefore, if we keep using $\epsilon = 0.1$ for Newton, the stopping condition may be too loose. We propose the following heuristic to change ϵ:

$$\begin{array}{ll} \epsilon \leftarrow 0.1\epsilon & \text{for logistic regression and L2-loss SVM} \\ \epsilon \leftarrow 0.001\epsilon & \text{for L2-loss SVR} \end{array}$$

The rationale is that if the default tolerance of dual CD has been used, we change ϵ to the default tolerance of primal Newton.

In Tables 3-10 we compare the primal Newton method and the **new** dual CD method. For the new dual CD, the rule of possibly switching to the primal Newton method is imposed and we consider two options.

- limit of CD iterations = 500
- limit of CD iterations = 300

The regularization parameters $C = C_{\text{best}}$ and $C = 100C_{\text{best}}$ are considered. We have the following observations.

- If $C = C_{\text{best}}$, the dual CD method is faster on document sets such as kdda, kddb, leisure.scale etc. However, if $C = 100C_{\text{best}}$, the primal Newton method is faster.

  This confirms the statements in Table 1.

- If the dual CD method reaches the iteration limit and primal Newton is called, we see that the number of Newton iterations needed is generally smaller than if the Newton method is directly run on the same problem. This result supports our choice of the initial $\boldsymbol{w}$ when switching to primal Newton.
- For some problems (e.g., kdda and kddb in Table 4), after switching from dual CD to primal Newton, the initial $\boldsymbol{w}$ already satisfies the stopping condition (1) and no Newton iteration is conducted. In such situations the default dual CD stopping condition may be too tight and slow convergence may not really occur. Having a stopping criterion neither too tight nor too loose is often difficult. The new setting of lowering the CD iteration limit and using primal Newton's stopping condition as a second check may help to avoid the over-solving of the optimization problem.
- If dual CD fails to meet the stopping condition after 300 iterations, then generally neither can it meet the condition after 500 iterations. Therefore, 200 CD iterations may be wasted without much progress. On the other hand, the needed Newton iterations after switching to Newton at the 300th or the 500th CD iteration are similar. Therefore, we decide to use 300 as the limit in the released code.

Because the default seed in the GNU C library is 1, for the same data, the five CV folds used in all experiments should be the same. However, we notice that for problem HIGGS, the second training procedure of dual CD takes

291 and 300

iterations, respectively in Tables 4 and 8. This result seems to be strange because the training set is the same. An explanation is as follows. In Table 4 dual CD runs 500 iterations for the first training procedure in the five-fold CV, while in Table 8, dual CD only runs 300. Thus for the next training procedure, the sequences of random numbers used are different. Note that dual CD randomly permutes all indices before each CD cycle. Thus the different numbers of CD iterations for the second training process are not an error.

3 Multi-core LIBLINEAR

The multi-core branch of LIBLINEAR has not been updated since version 2.30. We finished updating this branch and took this opportunity to conduct some experiments. We follow similar settings for Tables 7-10 to compare primal Newton and dual CD. Some details are given below.

- We run multi-core LIBLINEAR on the same computer and use 12 threads.
- Primal Newton involves some level-1 BLAS operations. Though they are not the bottleneck, we link OpenBLAS instead of compiling the BLAS code in LIBLINEAR. This may slightly reduce the running time.
- Parallel dual CD is available only for L2-loss SVM. Thus we do not report results of dual CD on logistic regression (i.e., solver `-s 7`).

Results are given in Tables 11-14 and we have the following observations.

- For logistic regression, in Table 7 with $C = C_{\text{best}}$, primal Newton is slower than dual CD in some situations. After using multi-core Newton, the running time is never slower than the standard single-thread dual CD.
- For L2-loss SVM, the speedup of multi-core primal Newton is better than multi-core dual CD. It is known that multi-core dual CD is more effective in the environment of a single CPU with multiple cores. The server we used has two CPUs so data access across CPUs causes a lower usage of the computational cores.

  In Table 13 of using $C = C_{\text{best}}$, while multi-core dual CD is still faster for some sets, the gap is smaller than that in Table 9. For example, in Table 9, for problem kdda, primal Newton is 10 times slower than dual CD. Now in Table 13, multi-core primal Newton is only three times slower than multi-core dual CD.

References

R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. LIBLINEAR: a library for large linear classification. *Journal of Machine Learning Research*, 9:1871–1874, 2008. URL `http://www.csie.ntu.edu.tw/~cjlin/papers/liblinear.pdf`.

Table 2: Results of running existing LIBLINEAR 2.41 on logistic regression. (`-s 7`: dual CD, `-s 0`: primal Newton) $C = 100C_{\text{best}}$.

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
HIGGS_100_s7	525	294	531	536	527	64.10	12886.59
HIGGS_100_s0	9	8	7	8	8	64.08	1486.03
a9a_100_s7	338	344	361	336	337	84.75	11.74
a9a_100_s0	5	5	5	5	6	84.77	0.61
astro-ph_62369.dat_100_s7	512	504	501	481	485	96.13	86.70
astro-ph_62369.dat_100_s0	7	7	7	7	7	96.27	11.24
australian_100_s7	1000	1000	1000	1000	1000	78.70	0.38
australian_100_s0	5	5	5	5	5	78.84	0.01
australian_scale_100_s7	13	12	11	11	11	85.80	0.01
australian_scale_100_s0	4	5	4	4	5	86.23	0.01
breast-cancer_100_s7	3	3	3	3	2	65.01	0.00
breast-cancer_100_s0	1	1	1	1	1	65.01	0.00
breast-cancer_scale_100_s7	1000	1000	1000	1000	1000	96.19	0.49
breast-cancer_scale_100_s0	5	5	5	6	5	96.78	0.01
cod-rna_100_s7	1000	1000	1000	1000	1000	89.08	54.77
cod-rna_100_s0	4	4	4	4	4	89.23	0.67
colon-cancer_100_s7	7	7	7	7	7	77.42	0.05
colon-cancer_100_s0	4	4	4	4	4	79.03	0.08
covtype.libsvm.binary.scale_100_s7	89	88	98	92	92	75.62	88.20
covtype.libsvm.binary.scale_100_s0	6	11	7	6	6	75.60	37.32
covtype.libsvm.binary_100_s7	1000	1000	1000	1000	1000	68.98	1004.92
covtype.libsvm.binary_100_s0	8	7	9	6	6	71.25	32.17
diabetes_100_s7	1000	1000	1000	1000	1000	67.97	0.42
diabetes_100_s0	3	2	3	3	3	67.84	0.01
diabetes_scale_100_s7	1000	1000	1000	1000	1000	76.69	0.42
diabetes_scale_100_s0	3	3	4	5	3	77.34	0.01
duke_100_s7	14	12	13	13	12	88.64	0.16
duke_100_s0	5	5	5	5	5	88.64	0.23
fourclass_100_s7	72	66	67	69	68	73.20	0.03
fourclass_100_s0	1	3	3	2	1	73.32	0.00
fourclass_scale_100_s7	6	6	6	8	7	68.68	0.01
fourclass_scale_100_s0	3	3	3	3	3	68.68	0.00
german.numer_100_s7	1000	1000	1000	1000	1000	71.60	0.77
german.numer_100_s0	4	4	4	4	4	76.60	0.02
german.numer_scale_100_s7	829	829	841	842	839	77.30	0.63
german.numer_scale_100_s0	4	4	4	4	4	77.00	0.02
gisette_scale_100_s7	66	70	68	66	70	97.15	34.15
gisette_scale_100_s0	6	7	7	6	7	97.15	31.30
heart_100_s7	1000	1000	1000	1000	1000	82.22	0.17
heart_100_s0	5	5	5	6	5	82.59	0.00
heart_scale_100_s7	7	7	7	7	7	82.96	0.00
heart_scale_100_s0	4	4	4	4	4	82.96	0.00
ijcnn1_100_s7	1000	1000	1000	1000	1000	91.21	75.06
ijcnn1_100_s0	7	6	6	7	7	92.46	1.51
ionosphere_scale_100_s7	1000	1000	1000	1000	1000	82.34	0.32

Table 2 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
ionosphere_scale_100_s0	6	6	5	5	5	84.62	0.01
kdda_100_s7	821	830	805	815	824	86.75	23562.43
kdda_100_s0	9	7	7	9	8	86.65	3539.57
kddb_100_s7	1000	1000	1000	1000	1000	87.82	69366.84
kddb_100_s0	7	7	7	10	7	87.81	5393.68
leisure.scale_100_s7	1000	1000	1000	1000	1000	85.75	5131.87
leisure.scale_100_s0	16	10	9	14	13	86.28	1587.37
leu_100_s7	9	10	9	9	10	94.74	0.13
leu_100_s0	6	6	6	6	6	92.11	0.22
liver-disorders_100_s7	486	508	577	539	557	68.28	0.04
liver-disorders_100_s0	4	4	3	3	3	68.97	0.00
liver-disorders_scale_100_s7	92	108	102	110	97	73.79	0.01
liver-disorders_scale_100_s0	4	4	3	3	3	73.79	0.00
madelon_100_s7	345	344	345	350	346	57.25	5.49
madelon_100_s0	6	5	5	6	6	57.20	0.91
mushrooms_100_s7	40	64	50	49	46	100.00	0.41
mushrooms_100_s0	6	6	6	6	6	99.98	0.16
news20.binary_100_s7	1000	1000	1000	1000	1000	96.65	329.45
news20.binary_100_s0	12	15	6	14	11	97.30	73.80
rcv1_test.binary_100_s7	373	364	369	376	376	97.16	954.74
rcv1_test.binary_100_s0	7	7	6	8	7	97.55	123.04
rcv1_train.binary_100_s7	1000	1000	1000	1000	1000	96.65	48.22
rcv1_train.binary_100_s0	6	6	6	6	6	96.54	2.99
real-sim_100_s7	668	675	672	679	674	97.12	109.85
real-sim_100_s0	7	7	7	7	7	97.50	8.41
skin_nonskin_100_s7	64	59	64	54	71	90.66	13.31
skin_nonskin_100_s0	4	4	4	4	4	90.66	3.43
sonar_scale_100_s7	1000	1000	1000	1000	1000	72.60	0.33
sonar_scale_100_s0	9	11	8	7	9	71.63	0.03
splice_100_s7	175	169	172	176	169	79.70	0.28
splice_100_s0	4	4	4	4	4	79.60	0.05
splice_scale_100_s7	32	32	32	32	32	71.00	0.08
splice_scale_100_s0	3	3	3	3	3	71.20	0.04
svmguide1_100_s7	169	172	181	175	174	83.52	0.34
svmguide1_100_s0	5	5	5	5	5	83.52	0.02
svmguide3_100_s7	1000	1000	1000	1000	1000	78.60	1.19
svmguide3_100_s0	6	4	6	7	4	79.73	0.04
train308.scale_100_s7	312	307	314	308	315	91.89	268.80
train308.scale_100_s0	7	9	7	7	7	92.06	84.72
url_combined_100_s7	84	85	84	86	85	99.20	1155.46
url_combined_100_s0	6	6	7	6	6	98.94	852.82
w8a_100_s7	1000	1000	1000	1000	1000	98.26	65.43
w8a_100_s0	8	9	10	10	10	98.36	1.18
webspam_wc_normalized_unigram.svm_100_s7	133	144	152	159	131	92.76	177.73
webspam_wc_normalized_unigram.svm_100_s0	6	7	7	7	7	92.73	61.26

Table 3: Logistic Regression (`-s 7`: dual CD, `-s 0`: primal Newton). $C = C_{\text{best}}$. For dual CD, if # iterations exceeds 500, primal Newton is called and the # of Newton iterations is shown in the next row.

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
HIGGS_s7	15	15	14	15	17	64.08	620.26
HIGGS_s0	7	8	7	7	6	64.06	1187.92
a9a_s7	14	16	15	14	11	84.78	0.54
a9a_s0	5	5	5	5	5	84.76	0.60
astro-ph_62369.dat_s7	10	10	11	10	9	96.82	3.85
astro-ph_62369.dat_s0	6	5	5	5	6	96.83	8.75
australian_s7	179	229	179	181	182	70.00	0.06
australian_s0	5	5	5	5	5	70.00	0.00
australian_scale_s7	6	5	5	6	5	86.96	0.00
australian_scale_s0	3	2	2	2	3	86.96	0.00
breast-cancer_s7	3	3	2	2	3	65.01	0.00
breast-cancer_s0	1	1	1	1	1	65.01	0.00
breast-cancer_scale_s7	500	500	500	500	457	96.78	0.18
	0	0	0	0	0		
breast-cancer_scale_s0	5	5	5	6	5	96.78	0.00
cod-rna_s7	79	63	78	67	71	87.58	4.25
cod-rna_s0	4	4	4	4	4	87.58	0.76
colon-cancer_s7	3	4	3	4	3	83.87	0.04
colon-cancer_s0	1	1	1	1	1	83.87	0.05
covtype.libsvm.binary.scale_s7	12	12	11	14	14	75.65	16.05
covtype.libsvm.binary.scale_s0	9	10	7	7	9	75.66	39.27
covtype.libsvm.binary_s7	218	204	220	188	203	61.62	189.14
covtype.libsvm.binary_s0	4	4	4	4	4	61.62	18.10
diabetes_s7	234	228	225	221	160	67.84	0.07
diabetes_s0	2	2	3	2	2	67.84	0.00
diabetes_scale_s7	69	70	74	75	61	77.34	0.03
diabetes_scale_s0	3	3	4	4	3	77.21	0.00
duke_s7	8	6	7	7	6	88.64	0.13
duke_s0	3	3	3	3	3	88.64	0.17
fourclass_s7	6	7	5	5	8	73.78	0.00
fourclass_s0	3	0	0	0	0	73.78	0.00
fourclass_scale_s7	4	4	4	4	4	68.68	0.00
fourclass_scale_s0	2	2	2	2	2	68.68	0.00
german.numer_s7	500	500	500	500	500	76.20	0.38
	1	1	1	1	1		
german.numer_s0	3	3	3	3	3	76.20	0.01
german.numer_scale_s7	15	15	13	17	16	77.00	0.02
german.numer_scale_s0	4	4	4	4	4	77.00	0.02
gisette_scale_s7	42	43	43	42	45	97.20	24.03
gisette_scale_s0	6	6	6	6	7	97.25	29.66
heart_s7	500	500	500	500	500	84.07	0.08
	3	3	3	3	4		
heart_s0	4	3	4	3	5	83.33	0.00
heart_scale_s7	4	3	3	4	4	83.33	0.00
heart_scale_s0	2	2	2	2	2	83.33	0.00
ijcnn1_s7	500	500	500	500	500	92.45	29.49

Table 3 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
	1	1	1	1	1		
ijcnn1_s0	7	6	6	7	7	92.45	1.51
ionosphere_scale_s7	500	500	500	500	500	84.33	0.14
	1	1	1	1	1		
ionosphere_scale_s0	6	6	6	6	5	84.33	0.01
kdda_s7	30	27	27	33	33	88.24	1026.16
kdda_s0	8	8	8	8	7	88.23	3158.82
kddb_s7	41	41	44	50	49	88.89	3004.30
kddb_s0	10	12	7	17	12	88.89	11303.21
leisure.scale_s7	75	76	76	77	78	87.25	416.21
leisure.scale_s0	8	10	9	10	8	87.27	760.93
leu_s7	7	9	10	8	8	89.47	0.11
leu_s0	6	6	6	6	6	92.11	0.20
liver-disorders_s7	9	10	10	9	10	70.34	0.00
liver-disorders_s0	3	3	3	2	2	70.34	0.00
liver-disorders_scale_s7	5	5	5	5	6	75.86	0.00
liver-disorders_scale_s0	2	3	2	2	2	75.86	0.00
madelon_s7	18	16	13	18	17	60.35	0.41
madelon_s0	1	1	2	1	2	60.40	0.32
mushrooms_s7	12	19	16	15	15	99.99	0.15
mushrooms_s0	6	6	6	6	6	99.98	0.16
news20.binary_s7	328	376	372	319	350	96.54	97.33
news20.binary_s0	6	6	6	6	6	96.79	22.15
rcv1_test.binary_s7	13	14	14	14	14	97.76	55.94
rcv1_test.binary_s0	7	7	7	7	7	97.74	102.89
rcv1_train.binary_s7	90	96	89	90	90	96.98	5.31
rcv1_train.binary_s0	6	6	6	6	6	97.02	2.84
real-sim_s7	14	16	13	12	12	97.53	3.96
real-sim_s0	6	6	6	6	6	97.53	6.59
skin_nonskin_s7	8	8	8	8	9	90.66	2.51
skin_nonskin_s0	4	4	4	4	4	90.66	3.35
sonar_scale_s7	39	42	43	45	45	74.04	0.01
sonar_scale_s0	4	5	4	6	4	74.04	0.01
splice_s7	9	12	7	8	7	80.80	0.02
splice_s0	3	2	3	2	3	80.70	0.02
splice_scale_s7	5	5	6	5	6	72.70	0.02
splice_scale_s0	2	2	2	2	2	72.70	0.02
svmguide1_s7	9	9	6	7	6	83.43	0.01
svmguide1_s0	5	5	5	5	5	83.46	0.01
svmguide3_s7	500	500	500	500	500	79.81	0.49
	3	4	3	3	3		
svmguide3_s0	5	5	4	6	4	79.57	0.02
train308.scale_s7	11	12	13	13	13	92.66	18.69
train308.scale_s0	6	7	8	7	7	92.66	64.80
url_combined_s7	13	12	11	12	13	98.25	211.41
url_combined_s0	5	5	6	5	7	98.09	601.90
w8a_s7	500	500	500	500	500	98.35	33.73

Table 3 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
	1	1	0	1	1		
w8a_s0	9	9	13	12	10	98.35	1.82
webspam_wc_normalized_unigram.svm_s7	9	13	11	12	12	92.33	24.58
webspam_wc_normalized_unigram.svm_s0	5	6	6	6	6	92.33	57.42

Table 4: Logistic Regression. (`-s 7`: dual CD, `-s 0`: primal Newton). $C = 100C_{\text{best}}$. For dual CD, if # iterations exceeds 500, primal Newton is called and the # of Newton iterations is shown in the next row.

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
HIGGS_100_s7	500	291	500	500	500	64.10	12859.18
	0	0	0	0	0		
HIGGS_100_s0	9	8	7	8	8	64.08	1325.56
a9a_100_s7	338	344	361	336	337	84.75	9.58
a9a_100_s0	5	5	5	5	6	84.77	0.53
astro-ph_62369.dat_100_s7	500	500	500	485	495	96.13	85.35
	0	0	0	0	0		
astro-ph_62369.dat_100_s0	7	7	7	7	7	96.27	10.78
australian_100_s7	500	500	500	500	500	78.70	0.21
	2	3	3	3	2		
australian_100_s0	5	5	5	5	5	78.84	0.00
australian_scale_100_s7	13	12	11	11	11	85.80	0.01
australian_scale_100_s0	4	5	4	4	5	86.23	0.00
breast-cancer_100_s7	3	3	3	3	2	65.01	0.00
breast-cancer_100_s0	1	1	1	1	1	65.01	0.00
breast-cancer_scale_100_s7	500	500	500	500	500	96.78	0.27
	2	2	3	2	3		
breast-cancer_scale_100_s0	5	5	5	6	5	96.78	0.00
cod-rna_100_s7	500	500	500	500	500	89.24	30.57
	1	2	2	2	2		
cod-rna_100_s0	4	4	4	4	4	89.23	0.80
colon-cancer_100_s7	7	7	7	7	7	77.42	0.05
colon-cancer_100_s0	4	4	4	4	4	79.03	0.07
covtype.libsvm.binary.scale_100_s7	89	88	98	92	92	75.62	92.68
covtype.libsvm.binary.scale_100_s0	6	11	7	6	6	75.60	31.97
covtype.libsvm.binary_100_s7	500	500	500	500	500	71.31	534.71
	7	7	7	5	5		
covtype.libsvm.binary_100_s0	8	7	9	6	6	71.25	31.38
diabetes_100_s7	500	500	500	500	500	67.97	0.21
	3	3	3	5	2		
diabetes_100_s0	3	2	3	3	3	67.84	0.00
diabetes_scale_100_s7	500	500	500	500	500	77.34	0.21
	4	3	4	4	5		
diabetes_scale_100_s0	3	3	4	5	3	77.34	0.00
duke_100_s7	14	12	13	13	12	88.64	0.14
duke_100_s0	5	5	5	5	5	88.64	0.21
fourclass_100_s7	72	66	67	69	68	73.20	0.02
fourclass_100_s0	1	3	3	2	1	73.32	0.00
fourclass_scale_100_s7	6	6	6	8	7	68.68	0.00
fourclass_scale_100_s0	3	3	3	3	3	68.68	0.00
german.numer_100_s7	500	500	500	500	500	76.60	0.40
	4	3	4	4	5		
german.numer_100_s0	4	4	4	4	4	76.60	0.01
german.numer_scale_100_s7	500	500	500	500	500	77.20	0.39
	1	1	1	2	2		
german.numer_scale_100_s0	4	4	4	4	4	77.00	0.01

Table 4 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
gisette_scale_100_s7	66	70	68	66	70	97.15	34.31
gisette_scale_100_s0	6	7	7	6	7	97.15	31.30
heart_100_s7	500	500	500	500	500	82.59	0.08
	4	5	4	6	9		
heart_100_s0	5	5	5	6	5	82.59	0.00
heart_scale_100_s7	7	7	7	7	7	82.96	0.00
heart_scale_100_s0	4	4	4	4	4	82.96	0.00
ijcnn1_100_s7	500	500	500	500	500	92.47	35.11
	4	4	6	6	6		
ijcnn1_100_s0	7	6	6	7	7	92.46	1.34
ionosphere_scale_100_s7	500	500	500	500	500	84.62	0.17
	9	5	6	6	5		
ionosphere_scale_100_s0	6	6	5	5	5	84.62	0.01
kdda_100_s7	500	500	500	500	500	86.75	14831.57
	0	0	0	0	0		
kdda_100_s0	9	7	7	9	8	86.65	3581.88
kddb_100_s7	500	500	500	500	500	87.82	34125.98
	0	0	0	0	0		
kddb_100_s0	7	7	7	10	7	87.81	4772.23
leisure.scale_100_s7	500	500	500	500	500	85.88	2802.63
	2	2	6	1	2		
leisure.scale_100_s0	16	10	9	14	13	86.28	1637.36
leu_100_s7	9	10	9	9	10	94.74	0.11
leu_100_s0	6	6	6	6	6	92.11	0.20
liver-disorders_100_s7	486	500	500	500	500	68.97	0.03
	0	1	0	1	1		
liver-disorders_100_s0	4	4	3	3	3	68.97	0.00
liver-disorders_scale_100_s7	92	108	102	110	97	73.79	0.01
liver-disorders_scale_100_s0	4	4	3	3	3	73.79	0.00
madelon_100_s7	345	344	345	350	346	57.25	5.03
madelon_100_s0	6	5	5	6	6	57.20	0.83
mushrooms_100_s7	40	64	50	49	46	100.00	0.35
mushrooms_100_s0	6	6	6	6	6	99.98	0.14
news20.binary_100_s7	500	500	500	500	500	96.69	164.55
	2	1	1	1	1		
news20.binary_100_s0	12	15	6	14	11	97.30	67.20
rcv1_test.binary_100_s7	373	364	369	376	376	97.16	1000.91
rcv1_test.binary_100_s0	7	7	6	8	7	97.55	124.18
rcv1_train.binary_100_s7	500	500	500	500	500	96.67	25.97
	1	0	0	0	0		
rcv1_train.binary_100_s0	6	6	6	6	6	96.54	2.74
real-sim_100_s7	500	500	500	500	500	97.12	87.00
	0	0	0	0	0		
real-sim_100_s0	7	7	7	7	7	97.50	7.70
skin_nonskin_100_s7	64	59	64	54	71	90.66	15.54
skin_nonskin_100_s0	4	4	4	4	4	90.66	3.33
sonar_scale_100_s7	500	500	500	500	500	72.12	0.14

Table 4 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
	3	3	3	5	4		
sonar_scale_100_s0	9	11	8	7	9	71.63	0.02
splice_100_s7	175	169	172	176	169	79.70	0.27
splice_100_s0	4	4	4	4	4	79.60	0.03
splice_scale_100_s7	32	32	32	32	32	71.00	0.05
splice_scale_100_s0	3	3	3	3	3	71.20	0.03
svmguide1_100_s7	169	172	181	175	174	83.52	0.27
svmguide1_100_s0	5	5	5	5	5	83.52	0.01
svmguide3_100_s7	500	500	500	500	500	79.89	0.50
	7	7	6	4	5		
svmguide3_100_s0	6	4	6	7	4	79.73	0.03
train308.scale_100_s7	312	307	314	308	315	91.89	271.37
train308.scale_100_s0	7	9	7	7	7	92.06	84.20
url_combined_100_s7	84	85	84	86	85	99.20	1046.04
url_combined_100_s0	6	6	7	6	6	98.94	757.98
w8a_100_s7	500	500	500	500	500	98.36	48.32
	5	5	5	8	7		
w8a_100_s0	8	9	10	10	10	98.36	1.80
webspam_wc_normalized_unigram.svm_100_s7	133	144	152	159	131	92.76	186.50
webspam_wc_normalized_unigram.svm_100_s0	6	7	7	7	7	92.73	66.26

Table 5: L2-loss SVM. (`-s 1`: dual CD, `-s 2`: primal Newton). $C = C_{\text{best}}$. For dual CD, if # iterations exceeds 500, primal Newton is called and the # of Newton iterations is shown in the next row.

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
HIGGS_s1	8	6	6	8	7	64.01	310.26
HIGGS_s2	9	8	8	7	8	64.00	1135.69
a9a_s1	14	13	14	13	14	84.80	0.29
a9a_s2	4	4	4	4	4	84.83	0.33
astro-ph_62369.dat_s1	8	8	8	8	8	96.94	2.88
astro-ph_62369.dat_s2	5	5	5	5	5	96.95	5.81
australian_s1	422	402	386	260	419	69.13	0.07
australian_s2	5	4	4	5	5	68.99	0.00
australian_scale_s1	4	4	4	4	4	86.67	0.00
australian_scale_s2	2	2	2	2	2	86.67	0.00
breast-cancer_s1	1	1	1	1	1	65.01	0.00
breast-cancer_s2	1	1	1	1	1	65.01	0.00
breast-cancer_scale_s1	25	22	26	25	27	96.78	0.00
breast-cancer_scale_s2	4	4	4	4	4	96.78	0.00
cod-rna_s1	53	54	52	47	53	87.57	1.53
cod-rna_s2	3	3	3	3	3	87.57	0.52
colon-cancer_s1	3	3	3	3	3	83.87	0.04
colon-cancer_s2	1	1	1	1	1	83.87	0.04
covtype.libsvm.binary.scale_s1	6	6	6	6	6	75.71	7.28
covtype.libsvm.binary.scale_s2	10	8	8	7	7	75.70	30.31
covtype.libsvm.binary_s1	22	23	23	23	23	61.25	19.81
covtype.libsvm.binary_s2	2	2	3	2	2	61.24	10.26
diabetes_s1	213	222	195	331	149	68.36	0.04
diabetes_s2	3	3	4	2	2	68.36	0.00
diabetes_scale_s1	9	9	9	9	9	77.34	0.00
diabetes_scale_s2	3	3	3	3	3	77.34	0.00
duke_s1	7	8	5	6	7	88.64	0.12
duke_s2	3	3	3	3	4	88.64	0.18
fourclass_s1	6	5	6	5	6	73.78	0.00
fourclass_s2	2	0	0	0	0	73.78	0.00
fourclass_scale_s1	4	4	4	4	5	68.68	0.00
fourclass_scale_s2	1	1	1	1	1	68.68	0.00
german.numer_s1	500	500	500	500	500	76.30	0.22
	2	1	1	2	1		
german.numer_s2	3	3	3	3	3	76.30	0.01
german.numer_scale_s1	80	80	79	80	80	76.80	0.03
german.numer_scale_s2	4	4	4	3	4	76.60	0.01
gisette_scale_s1	24	25	24	24	25	97.40	11.71
gisette_scale_s2	5	5	6	5	6	97.47	19.02
heart_s1	500	500	500	500	500	83.33	0.04
	3	4	3	2	2		
heart_s2	3	3	3	3	3	83.33	0.00
heart_scale_s1	3	3	3	3	4	83.33	0.00
heart_scale_s2	2	2	2	2	2	83.33	0.00
ijcnn1_s1	16	16	15	15	16	92.27	0.58

Table 5 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
ijcnn1_s2	6	6	6	6	5	92.25	0.73
ionosphere_scale_s1	311	394	436	317	326	84.90	0.04
ionosphere_scale_s2	5	6	5	5	5	84.90	0.01
kdda_s1	10	10	10	10	10	88.25	333.81
kdda_s2	16	14	15	14	16	88.24	3593.42
kddb_s1	25	24	24	24	24	88.95	1445.05
kddb_s2	17	15	15	15	17	88.95	7601.96
leisure.scale_s1	39	39	39	38	39	87.36	178.61
leisure.scale_s2	9	10	11	10	11	87.37	530.09
leu_s1	7	10	9	7	7	94.74	0.11
leu_s2	6	5	6	5	7	94.74	0.22
liver-disorders_s1	11	11	10	11	12	69.66	0.00
liver-disorders_s2	3	2	2	2	2	69.66	0.00
liver-disorders_scale_s1	6	6	6	6	6	75.86	0.00
liver-disorders_scale_s2	2	2	2	2	2	75.86	0.00
madelon_s1	9	9	9	9	9	59.95	0.31
madelon_s2	1	1	1	1	1	60.30	0.30
mushrooms_s1	36	65	30	36	43	100.00	0.09
mushrooms_s2	5	5	5	5	5	99.96	0.08
news20.binary_s1	330	357	345	335	356	96.77	48.12
news20.binary_s2	6	5	6	5	4	97.16	12.78
rcv1_test.binary_s1	10	10	10	10	10	97.81	32.30
rcv1_test.binary_s2	5	6	5	5	5	97.80	63.66
rcv1_train.binary_s1	11	11	11	11	10	97.08	0.91
rcv1_train.binary_s2	5	4	5	6	4	97.05	1.65
real-sim_s1	9	9	8	9	9	97.55	2.27
real-sim_s2	5	5	5	6	5	97.55	4.28
skin_nonskin_s1	7	7	7	7	7	90.83	1.54
skin_nonskin_s2	1	2	2	2	2	90.84	1.79
sonar_scale_s1	15	14	17	15	16	74.04	0.00
sonar_scale_s2	5	4	3	5	4	74.04	0.01
splice_s1	13	13	13	13	13	80.90	0.02
splice_s2	3	3	3	3	3	81.00	0.02
splice_scale_s1	5	5	5	5	5	72.60	0.02
splice_scale_s2	2	2	2	2	2	72.70	0.02
svmguide1_s1	6	6	6	6	6	83.36	0.01
svmguide1_s2	5	4	5	5	4	83.36	0.01
svmguide3_s1	500 2	500 1	500 3	500 2	500 1	79.57	0.28
svmguide3_s2	8	6	5	4	6	79.49	0.03
train308.scale_s1	9	9	9	9	9	92.81	13.13
train308.scale_s2	10	10	8	9	10	92.79	51.90
url_combined_s1	7	7	7	7	7	98.53	119.46
url_combined_s2	5	8	5	5	5	98.26	381.91
w8a_s1	500 1	500 2	500 3	500 1	500 1	98.22	3.09
w8a_s2	8	8	8	8	11	98.23	0.71

Table 5 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
webspam_wc_normalized_unigram.svm_s1	10	10	10	10	10	92.60	19.54
webspam_wc_normalized_unigram.svm_s2	5	5	4	4	5	92.57	41.67

Table 6: L2-loss SVM. (`-s 1`: dual CD, `-s 2`: primal Newton). $C = 100C_{\text{best}}$. For dual CD, if # iterations exceeds 500, primal Newton is called and the # of Newton iterations is shown in the next row.

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
HIGGS_100_s1	333	167	339	343	343	64.05	8043.69
HIGGS_100_s2	9	9	7	7	7	64.03	1211.18
a9a_100_s1	500	500	500	500	500	84.80	6.42
	1	1	1	1	1		
a9a_100_s2	4	4	4	4	4	84.75	0.33
astro-ph_62369.dat_100_s1	229	233	241	237	229	95.49	9.51
astro-ph_62369.dat_100_s2	5	5	6	5	5	95.49	5.78
australian_100_s1	500	500	500	500	500	79.57	0.10
	5	4	2	3	4		
australian_100_s2	4	5	4	5	4	79.57	0.00
australian_scale_100_s1	9	9	9	9	9	86.09	0.00
australian_scale_100_s2	4	5	5	4	4	86.09	0.00
breast-cancer_100_s1	5	3	3	5	4	65.01	0.00
breast-cancer_100_s2	2	2	2	2	1	65.01	0.00
breast-cancer_scale_100_s1	500	500	500	500	500	96.78	0.03
	1	1	1	2	2		
breast-cancer_scale_100_s2	4	5	4	6	4	96.78	0.00
cod-rna_100_s1	500	500	500	500	500	88.39	14.03
	1	3	2	2	3		
cod-rna_100_s2	3	3	3	3	3	88.40	0.55
colon-cancer_100_s1	8	9	7	8	9	75.81	0.05
colon-cancer_100_s2	4	4	4	4	5	75.81	0.07
covtype.libsvm.binary.scale_100_s1	46	45	45	44	44	75.68	33.11
covtype.libsvm.binary.scale_100_s2	9	10	10	9	9	75.69	33.04
covtype.libsvm.binary_100_s1	500	500	500	500	500	64.91	393.37
	2	3	4	2	1		
covtype.libsvm.binary_100_s2	5	8	6	7	6	64.95	21.54
diabetes_100_s1	500	500	500	500	500	67.45	0.09
	3	2	2	2	3		
diabetes_100_s2	1	3	5	2	4	67.45	0.00
diabetes_scale_100_s1	500	482	399	429	405	77.21	0.07
	1	0	0	0	0		
diabetes_scale_100_s2	3	3	3	3	3	77.34	0.00
duke_100_s1	11	13	13	13	15	88.64	0.14
duke_100_s2	9	6	7	8	4	90.91	0.22
fourclass_100_s1	75	70	72	74	71	73.32	0.01
fourclass_100_s2	2	2	2	2	2	73.32	0.00
fourclass_scale_100_s1	9	7	9	8	8	68.56	0.00
fourclass_scale_100_s2	2	2	2	2	2	68.45	0.00
german.numer_100_s1	500	500	500	500	500	76.70	0.24
	5	5	4	3	3		
german.numer_100_s2	3	3	4	3	3	76.80	0.01
german.numer_scale_100_s1	500	500	500	500	500	76.70	0.22
	4	4	4	4	4		
german.numer_scale_100_s2	3	4	3	3	4	76.90	0.01
gisette_scale_100_s1	62	66	62	64	66	96.97	15.68

Table 6 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
gisette_scale_100_s2	5	7	5	5	6	97.07	17.73
heart_100_s1	500	500	500	500	500	83.70	0.04
	4	4	4	5	4		
heart_100_s2	4	5	5	4	6	83.33	0.00
heart_scale_100_s1	8	8	8	9	9	82.96	0.00
heart_scale_100_s2	3	3	3	4	3	82.96	0.00
ijcnn1_100_s1	500	500	500	500	500	92.29	11.11
	2	2	2	2	2		
ijcnn1_100_s2	5	5	6	6	6	92.25	0.74
ionosphere_scale_100_s1	500	500	500	500	500	84.33	0.05
	7	7	8	5	5		
ionosphere_scale_100_s2	6	6	5	6	6	83.76	0.01
kdda_100_s1	401	423	407	409	421	86.52	5841.76
kdda_100_s2	11	11	13	13	14	86.49	2814.50
kddb_100_s1	500	500	500	500	500	87.49	20152.69
	1	1	1	1	1		
kddb_100_s2	17	18	18	18	18	87.52	10285.00
leisure.scale_100_s1	500	500	500	500	500	84.95	1486.11
	1	3	2	2	2		
leisure.scale_100_s2	6	7	6	7	9	85.96	329.70
leu_100_s1	7	8	9	9	8	94.74	0.11
leu_100_s2	4	4	5	4	6	97.37	0.14
liver-disorders_100_s1	500	500	500	500	500	68.97	0.01
	1	1	1	1	0		
liver-disorders_100_s2	3	1	2	1	1	68.97	0.00
liver-disorders_scale_100_s1	99	138	153	104	104	73.79	0.00
liver-disorders_scale_100_s2	3	3	2	3	3	74.48	0.00
madelon_100_s1	500	500	500	500	361	57.05	6.28
	1	1	1	1	0		
madelon_100_s2	7	4	4	6	5	57.20	0.80
mushrooms_100_s1	43	72	39	44	59	100.00	0.09
mushrooms_100_s2	5	5	5	5	6	99.98	0.07
news20.binary_100_s1	500	500	500	500	500	96.76	78.27
	1	2	2	1	1		
news20.binary_100_s2	5	5	6	5	4	97.25	10.15
rcv1_test.binary_100_s1	452	420	469	427	431	97.00	203.21
rcv1_test.binary_100_s2	7	10	9	6	7	97.29	70.47
rcv1_train.binary_100_s1	500	500	500	500	500	96.32	5.44
	0	0	0	0	0		
rcv1_train.binary_100_s2	4	6	5	5	5	96.06	1.51
real-sim_100_s1	351	347	347	350	375	96.65	7.74
real-sim_100_s2	7	6	6	6	6	96.74	4.25
skin_nonskin_100_s1	302	298	285	206	290	90.83	44.68
skin_nonskin_100_s2	2	2	1	1	2	90.84	1.78
sonar_scale_100_s1	500	500	500	500	500	74.52	0.05
	2	2	2	3	2		
sonar_scale_100_s2	7	8	10	8	9	73.08	0.00
splice_100_s1	500	500	500	500	500	80.10	0.34

Table 6 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
	2	1	1	1	1		
splice_100_s2	4	4	4	4	4	80.10	0.03
splice_scale_100_s1	36	33	34	33	35	71.00	0.04
splice_scale_100_s2	3	3	3	3	3	71.00	0.02
svmguide1_100_s1	95	100	94	94	93	83.49	0.04
svmguide1_100_s2	5	5	5	5	5	83.46	0.01
svmguide3_100_s1	500	500	500	500	500	79.49	0.29
	4	6	3	5	5		
svmguide3_100_s2	7	6	5	6	4	79.32	0.02
train308.scale_100_s1	378	371	379	376	375	91.08	104.81
train308.scale_100_s2	11	10	10	10	11	91.56	48.93
url_combined_100_s1	180	180	180	181	168	99.44	487.37
url_combined_100_s2	9	7	6	7	5	99.21	409.74
w8a_100_s1	500	500	500	500	500	98.21	3.84
	6	7	7	6	5		
w8a_100_s2	12	11	9	9	9	98.23	0.73
webspam_wc_normalized_unigram.svm_100_s1	477	470	500	481	500	92.71	381.08
	0	0	1	0	0		
webspam_wc_normalized_unigram.svm_100_s2	5	5	4	5	5	92.66	43.39

Table 7: Logistic Regression (`-s 7`: dual CD, `-s 0`: primal Newton). $C = C_{\text{best}}$. For dual CD, if # iterations exceeds 300, primal Newton is called and the # of Newton iterations is shown in the next row.

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
HIGGS_s7	15	15	14	15	17	64.08	566.54
HIGGS_s0	7	8	7	7	6	64.06	1204.51
a9a_s7	14	16	15	14	11	84.78	0.54
a9a_s0	5	5	5	5	5	84.76	0.51
astro-ph_62369.dat_s7	10	10	11	10	9	96.82	3.91
astro-ph_62369.dat_s0	6	5	5	5	6	96.83	9.47
australian_s7	179	229	179	181	182	70.00	0.06
australian_s0	5	5	5	5	5	70.00	0.04
australian_scale_s7	6	5	5	6	5	86.96	0.01
australian_scale_s0	3	2	2	2	3	86.96	0.01
breast-cancer_s7	3	3	2	2	3	65.01	0.00
breast-cancer_s0	1	1	1	1	1	39.09	0.00
breast-cancer_scale_s7	300	300	300	300	300	96.78	0.13
	0	0	0	0	0		
breast-cancer_scale_s0	5	5	5	6	5	96.78	0.01
cod-rna_s7	79	63	78	67	71	87.58	3.03
cod-rna_s0	4	4	4	4	4	87.58	0.58
colon-cancer_s7	3	4	3	4	3	83.87	0.05
colon-cancer_s0	1	1	1	1	1	83.87	0.24
covtype.libsvm.binary.scale_s7	12	12	11	14	14	75.65	15.76
covtype.libsvm.binary.scale_s0	9	10	7	7	9	75.66	38.84
covtype.libsvm.binary_s7	218	204	220	188	203	61.62	192.28
covtype.libsvm.binary_s0	4	4	4	4	4	61.62	18.77
diabetes_s7	234	228	225	221	160	67.84	0.08
diabetes_s0	2	2	3	2	2	67.84	0.01
diabetes_scale_s7	69	70	74	75	61	77.34	0.03
diabetes_scale_s0	3	3	4	4	3	77.21	0.01
duke_s7	8	6	7	7	6	88.64	0.13
duke_s0	3	3	3	3	3	88.64	0.28
fourclass_s7	6	7	5	5	8	73.78	0.00
fourclass_s0	3	3	3	3	3	73.78	0.00
fourclass_scale_s7	4	4	4	4	4	68.68	0.00
fourclass_scale_s0	2	2	2	2	2	68.68	0.00
german.numer_s7	300	300	300	300	300	76.10	0.24
	3	2	1	2	1		
german.numer_s0	3	3	3	3	3	76.20	0.03
german.numer_scale_s7	15	15	13	17	16	77.00	0.02
german.numer_scale_s0	4	4	4	4	4	77.00	0.03
gisette_scale_s7	42	43	43	42	45	97.20	24.34
gisette_scale_s0	6	6	6	6	7	97.25	33.02
heart_s7	300	300	300	300	300	83.33	0.05
	4	3	4	4	4		
heart_s0	4	3	4	3	5	83.33	0.02
heart_scale_s7	4	3	3	4	4	83.33	0.00
heart_scale_s0	2	2	2	2	2	83.33	0.00
ijcnn1_s7	300	300	300	300	300	92.46	14.90

Table 7 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
	2	2	2	2	2		
ijcnn1_s0	7	6	6	7	7	92.45	0.95
ionosphere_scale_s7	300	300	300	300	300	84.33	0.09
	2	1	2	2	1		
ionosphere_scale_s0	6	6	6	6	5	84.33	0.01
kdda_s7	30	27	27	33	33	88.24	998.50
kdda_s0	8	8	8	8	7	88.23	3171.60
kddb_s7	41	41	44	50	49	88.89	3090.66
kddb_s0	10	12	7	17	12	88.89	11578.24
leisure.scale_s7	75	76	76	77	78	87.25	419.46
leisure.scale_s0	8	10	9	10	8	87.27	758.93
leu_s7	7	9	10	8	8	89.47	0.12
leu_s0	6	6	6	6	6	92.11	0.22
liver-disorders_s7	9	10	10	9	10	70.34	0.00
liver-disorders_s0	3	3	3	2	2	70.34	0.00
liver-disorders_scale_s7	5	5	5	5	6	75.86	0.00
liver-disorders_scale_s0	2	3	2	2	2	75.86	0.00
madelon_s7	18	16	13	18	17	60.35	0.42
madelon_s0	1	1	2	1	2	60.40	0.35
mushrooms_s7	12	19	16	15	15	99.99	0.16
mushrooms_s0	6	6	6	6	6	99.98	0.17
news20.binary_s7	300	300	300	300	300	96.54	85.82
	0	0	0	0	0		
news20.binary_s0	6	6	6	6	6	96.79	23.62
rcv1_test.binary_s7	13	14	14	14	14	97.76	60.90
rcv1_test.binary_s0	7	7	7	7	7	97.74	107.28
rcv1_train.binary_s7	90	96	89	90	90	96.98	4.69
rcv1_train.binary_s0	6	6	6	6	6	97.02	2.75
real-sim_s7	14	16	13	12	12	97.53	3.98
real-sim_s0	6	6	6	6	6	97.53	6.67
skin_nonskin_s7	8	8	8	8	9	90.66	2.28
skin_nonskin_s0	4	4	4	4	4	90.66	2.99
sonar_scale_s7	39	42	43	45	45	74.04	0.02
sonar_scale_s0	4	5	4	6	4	74.04	0.01
splice_s7	9	12	7	8	7	80.80	0.03
splice_s0	3	2	3	2	3	80.70	0.03
splice_scale_s7	5	5	6	5	6	72.70	0.02
splice_scale_s0	2	2	2	2	2	72.70	0.03
svmguide1_s7	9	9	6	7	6	83.43	0.02
svmguide1_s0	5	5	5	5	5	83.46	0.02
svmguide3_s7	300	300	300	300	300	79.81	0.32
	3	4	3	4	3		
svmguide3_s0	5	5	4	6	4	79.57	0.03
train308.scale_s7	11	12	13	13	13	92.66	20.43
train308.scale_s0	6	7	8	7	7	92.66	67.76
url_combined_s7	13	12	11	12	13	98.25	214.91
url_combined_s0	5	5	6	5	7	98.09	612.64

Table 7 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
w8a_s7	300	300	300	300	300	98.36	15.38
	2	1	1	0	2		
w8a_s0	9	9	13	12	10	98.35	1.40
webspam_wc_normalized_unigram.svm_s7	9	13	11	12	12	92.33	28.44
webspam_wc_normalized_unigram.svm_s0	5	6	6	6	6	92.33	56.16

Table 8: Logistic Regression. (`-s 7`: dual CD, `-s 0`: primal Newton). $C = 100C_{\text{best}}$. For dual CD, if # iterations exceeds 300, primal Newton is called and the # of Newton iterations is shown in the next row.

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
HIGGS_100_s7	300	300	300	300	300	64.10	8095.74
	0	0	0	0	0		
HIGGS_100_s0	9	8	7	8	8	64.08	1214.53
a9a_100_s7	300	300	300	300	300	84.79	8.58
	1	0	1	1	1		
a9a_100_s0	5	5	5	5	6	84.77	0.49
astro-ph_62369.dat_100_s7	300	300	300	300	300	96.13	51.02
	0	0	0	0	0		
astro-ph_62369.dat_100_s0	7	7	7	7	7	96.27	10.60
australian_100_s7	300	300	300	300	300	78.84	0.13
	2	2	3	2	2		
australian_100_s0	5	5	5	5	5	78.84	0.01
australian_scale_100_s7	13	12	11	11	11	85.80	0.01
australian_scale_100_s0	4	5	4	4	5	86.23	0.01
breast-cancer_100_s7	3	3	3	3	2	65.01	0.00
breast-cancer_100_s0	1	1	1	1	1	65.01	0.00
breast-cancer_scale_100_s7	300	300	300	300	300	96.63	0.17
	2	3	2	3	2		
breast-cancer_scale_100_s0	5	5	5	6	5	96.78	0.01
cod-rna_100_s7	300	300	300	300	300	89.24	13.53
	3	2	2	1	2		
cod-rna_100_s0	4	4	4	4	4	89.23	0.52
colon-cancer_100_s7	7	7	7	7	7	77.42	0.05
colon-cancer_100_s0	4	4	4	4	4	79.03	0.08
covtype.libsvm.binary.scale_100_s7	89	88	98	92	92	75.62	89.45
covtype.libsvm.binary.scale_100_s0	6	11	7	6	6	75.60	31.30
covtype.libsvm.binary_100_s7	300	300	300	300	300	71.31	321.39
	6	5	7	9	7		
covtype.libsvm.binary_100_s0	8	7	9	6	6	71.25	31.41
diabetes_100_s7	300	300	300	300	300	68.10	0.13
	7	4	3	4	5		
diabetes_100_s0	3	2	3	3	3	67.84	0.00
diabetes_scale_100_s7	300	300	300	300	300	77.34	0.13
	4	4	3	4	4		
diabetes_scale_100_s0	3	3	4	5	3	77.34	0.00
duke_100_s7	14	12	13	13	12	88.64	0.15
duke_100_s0	5	5	5	5	5	88.64	0.22
fourclass_100_s7	72	66	67	69	68	73.20	0.03
fourclass_100_s0	3	3	3	3	3	73.32	0.00
fourclass_scale_100_s7	6	6	6	8	7	68.68	0.00
fourclass_scale_100_s0	3	3	3	3	3	68.68	0.00
german.numer_100_s7	300	300	300	300	300	76.80	0.25
	4	4	5	3	4		
german.numer_100_s0	4	4	4	4	4	76.60	0.02
german.numer_scale_100_s7	300	300	300	300	300	77.00	0.25
	2	2	3	2	2		

Table 8 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
german.numer_scale_100_s0	4	4	4	4	4	77.00	0.02
gisette_scale_100_s7	66	70	68	66	70	97.15	34.66
gisette_scale_100_s0	6	7	7	6	7	97.15	31.56
heart_100_s7	300	300	300	300	300	82.96	0.05
	5	5	8	5	5		
heart_100_s0	5	5	5	6	5	82.59	0.00
heart_scale_100_s7	7	7	7	7	7	82.96	0.00
heart_scale_100_s0	4	4	4	4	4	82.96	0.00
ijcnn1_100_s7	300	300	300	300	300	92.46	17.84
	4	6	5	5	5		
ijcnn1_100_s0	7	6	6	7	7	92.46	0.95
ionosphere_scale_100_s7	300	300	300	300	300	84.62	0.12
	6	7	6	10	7		
ionosphere_scale_100_s0	6	6	5	5	5	84.62	0.01
kdda_100_s7	300	300	300	300	300	86.75	8603.58
	1	0	1	0	0		
kdda_100_s0	9	7	7	9	8	86.65	3576.87
kddb_100_s7	300	300	300	300	300	87.82	20468.49
	0	0	0	0	0		
kddb_100_s0	7	7	7	10	7	87.81	4889.11
leisure.scale_100_s7	300	300	300	300	300	85.92	1873.69
	4	3	2	5	4		
leisure.scale_100_s0	16	10	9	14	13	86.28	1697.93
leu_100_s7	9	10	9	9	10	94.74	0.13
leu_100_s0	6	6	6	6	6	92.11	0.21
liver-disorders_100_s7	300	300	300	300	300	68.97	0.02
	1	1	1	0	1		
liver-disorders_100_s0	4	4	3	3	3	68.97	0.00
liver-disorders_scale_100_s7	92	108	102	110	97	73.79	0.01
liver-disorders_scale_100_s0	4	4	3	3	3	73.79	0.00
madelon_100_s7	300	300	300	300	300	57.20	4.28
	1	1	1	2	1		
madelon_100_s0	6	5	5	6	6	57.20	0.81
mushrooms_100_s7	40	64	50	49	46	100.00	0.36
mushrooms_100_s0	6	6	6	6	6	99.98	0.15
news20.binary_100_s7	300	300	300	300	300	96.66	103.35
	2	6	2	2	2		
news20.binary_100_s0	12	15	6	14	11	97.30	68.51
rcv1_test.binary_100_s7	300	300	300	300	300	97.17	838.74
	0	0	0	0	0		
rcv1_test.binary_100_s0	7	7	6	8	7	97.55	123.54
rcv1_train.binary_100_s7	300	300	300	300	300	96.72	16.29
	1	0	0	0	0		
rcv1_train.binary_100_s0	6	6	6	6	6	96.54	2.67
real-sim_100_s7	300	300	300	300	300	97.12	53.99
	0	0	0	0	0		
real-sim_100_s0	7	7	7	7	7	97.50	7.89

Table 8 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
skin_nonskin_100_s7	64	59	64	54	71	90.66	13.50
skin_nonskin_100_s0	4	4	4	4	4	90.66	2.73
sonar_scale_100_s7	300	300	300	300	300	72.12	0.10
	6	5	5	5	6		
sonar_scale_100_s0	9	11	8	7	9	71.63	0.02
splice_100_s7	175	169	172	176	169	79.70	0.24
splice_100_s0	4	4	4	4	4	79.60	0.04
splice_scale_100_s7	32	32	32	32	32	71.00	0.06
splice_scale_100_s0	3	3	3	3	3	71.20	0.03
svmguide1_100_s7	169	172	181	175	174	83.52	0.29
svmguide1_100_s0	5	5	5	5	5	83.52	0.02
svmguide3_100_s7	300	300	300	300	300	79.89	0.35
	8	6	6	5	6		
svmguide3_100_s0	6	4	6	7	4	79.73	0.03
train308.scale_100_s7	300	300	300	300	300	91.90	274.52
	0	0	0	0	0		
train308.scale_100_s0	7	9	7	7	7	92.06	85.03
url_combined_100_s7	84	85	84	86	85	99.20	1029.32
url_combined_100_s0	6	6	7	6	6	98.94	780.83
w8a_100_s7	300	300	300	300	300	98.37	23.19
	9	9	6	8	6		
w8a_100_s0	8	9	10	10	10	98.36	1.18
webspam_wc_normalized_unigram.svm_100_s7	133	144	152	159	131	92.76	188.84
webspam_wc_normalized_unigram.svm_100_s0	6	7	7	7	7	92.73	65.65

Table 9: L2-loss SVM. (`-s 1`: dual CD, `-s 2`: primal Newton). $C = C_{\text{best}}$. For dual CD, if # iterations exceeds 300, primal Newton is called and the # of Newton iterations is shown in the next row.

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
HIGGS_s1	8	6	6	8	7	64.01	320.32
HIGGS_s2	9	8	8	7	8	64.00	1137.87
a9a_s1	14	13	14	13	14	84.80	0.30
a9a_s2	4	4	4	4	4	84.83	0.34
astro-ph_62369.dat_s1	8	8	8	8	8	96.94	2.90
astro-ph_62369.dat_s2	5	5	5	5	5	96.95	5.89
australian_s1	300	300	300	255	300	68.99	0.06
	1	1	0	0	1		
australian_s2	5	4	4	5	5	68.99	0.01
australian_scale_s1	4	4	4	4	4	86.67	0.00
australian_scale_s2	2	2	2	2	2	86.67	0.00
breast-cancer_s1	2	2	2	2	2	65.01	0.00
breast-cancer_s2	1	1	1	1	1	39.09	0.00
breast-cancer_scale_s1	25	22	26	25	27	96.78	0.00
breast-cancer_scale_s2	4	4	4	4	4	96.78	0.00
cod-rna_s1	53	54	52	47	53	87.57	1.12
cod-rna_s2	3	3	3	3	3	87.57	0.37
colon-cancer_s1	3	3	3	3	3	83.87	0.05
colon-cancer_s2	1	1	1	1	1	83.87	0.05
covtype.libsvm.binary.scale_s1	6	6	6	6	6	75.71	7.18
covtype.libsvm.binary.scale_s2	10	8	8	7	7	75.70	30.22
covtype.libsvm.binary_s1	22	23	23	23	23	61.25	18.46
covtype.libsvm.binary_s2	2	2	3	2	2	61.24	10.35
diabetes_s1	213	222	195	300	291	68.36	0.04
	0	0	0	1	0		
diabetes_s2	3	3	4	2	2	68.36	0.00
diabetes_scale_s1	9	9	9	9	9	77.34	0.00
diabetes_scale_s2	3	3	3	3	3	77.34	0.00
duke_s1	7	8	5	6	7	88.64	0.13
duke_s2	3	3	3	3	4	88.64	0.19
fourclass_s1	6	5	6	5	6	73.78	0.00
fourclass_s2	2	2	2	2	2	73.78	0.00
fourclass_scale_s1	4	4	4	4	5	68.68	0.00
fourclass_scale_s2	1	1	1	1	1	68.68	0.00
german.numer_s1	300	300	300	300	300	76.30	0.14
	2	2	2	2	2		
german.numer_s2	3	3	3	3	3	76.30	0.01
german.numer_scale_s1	80	80	79	80	80	76.80	0.04
german.numer_scale_s2	4	4	4	3	4	76.60	0.01
gisette_scale_s1	24	25	24	24	25	97.40	12.36
gisette_scale_s2	5	5	6	5	6	97.47	19.42
heart_s1	300	300	300	300	300	83.33	0.03
	3	2	3	3	2		
heart_s2	3	3	3	3	3	83.33	0.00
heart_scale_s1	3	3	3	3	4	83.33	0.00

Table 9 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
heart_scale_s2	2	2	2	2	2	83.33	0.00
ijcnn1_s1	16	16	15	15	16	92.27	0.52
ijcnn1_s2	6	6	6	6	5	92.25	0.85
ionosphere_scale_s1	300	300	300	291	300	84.33	0.04
	0	1	1	0	1		
ionosphere_scale_s2	5	6	5	5	5	84.90	0.01
kdda_s1	10	10	10	10	10	88.25	330.97
kdda_s2	16	14	15	14	16	88.24	3405.58
kddb_s1	25	24	24	24	24	88.95	1351.98
kddb_s2	17	15	15	15	17	88.95	7347.57
leisure.scale_s1	39	39	39	38	39	87.36	191.56
leisure.scale_s2	9	10	11	10	11	87.37	567.57
leu_s1	7	10	9	7	7	94.74	0.12
leu_s2	6	5	6	5	7	94.74	0.23
liver-disorders_s1	11	11	10	11	12	69.66	0.00
liver-disorders_s2	3	2	2	2	2	69.66	0.00
liver-disorders_scale_s1	6	6	6	6	6	75.86	0.00
liver-disorders_scale_s2	2	2	2	2	2	75.86	0.00
madelon_s1	9	9	9	9	9	59.95	0.34
madelon_s2	1	1	1	1	1	60.30	0.32
mushrooms_s1	36	65	30	36	43	100.00	0.10
mushrooms_s2	5	5	5	5	5	99.96	0.09
news20.binary_s1	300	300	300	300	300	96.77	43.68
	0	0	0	0	0		
news20.binary_s2	6	5	6	5	4	97.16	13.27
rcv1_test.binary_s1	10	10	10	10	10	97.81	36.89
rcv1_test.binary_s2	5	6	5	5	5	97.80	69.38
rcv1_train.binary_s1	11	11	11	11	10	97.08	0.93
rcv1_train.binary_s2	5	4	5	6	4	97.05	1.61
real-sim_s1	9	9	8	9	9	97.55	2.40
real-sim_s2	5	5	5	6	5	97.55	4.33
skin_nonskin_s1	7	7	7	7	7	90.83	1.27
skin_nonskin_s2	2	2	2	2	2	90.84	1.47
sonar_scale_s1	15	14	17	15	16	74.04	0.01
sonar_scale_s2	5	4	3	5	4	74.04	0.01
splice_s1	13	13	13	13	13	80.90	0.03
splice_s2	3	3	3	3	3	81.00	0.03
splice_scale_s1	5	5	5	5	5	72.60	0.02
splice_scale_s2	2	2	2	2	2	72.70	0.02
svmguide1_s1	6	6	6	6	6	83.36	0.01
svmguide1_s2	5	4	5	5	4	83.36	0.01
svmguide3_s1	300	300	300	300	300	79.49	0.18
	2	2	4	3	3		
svmguide3_s2	8	6	5	4	6	79.49	0.03
train308.scale_s1	9	9	9	9	9	92.81	13.91
train308.scale_s2	10	10	8	9	10	92.79	52.88
url_combined_s1	7	7	7	7	7	98.53	122.18

Table 9 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
url_combined_s2	5	8	5	5	5	98.26	385.17
w8a_s1	300	300	300	300	300	98.22	2.01
	3	2	3	2	2		
w8a_s2	8	8	8	8	11	98.23	0.54
webspam_wc_normalized_unigram.svm_s1	10	10	10	10	10	92.60	20.43
webspam_wc_normalized_unigram.svm_s2	5	5	4	4	5	92.57	44.03

Table 10: L2-loss SVM. (`-s 1`: dual CD, `-s 2`: primal Newton). $C = 100C_{\text{best}}$. For dual CD, if # iterations exceeds 300, primal Newton is called and the # of Newton iterations is shown in the next row.

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
HIGGS_100_s1	300	164	300	300	300	64.05	6784.00
	0	0	0	0	0		
HIGGS_100_s2	9	9	7	7	7	64.03	1225.11
a9a_100_s1	300	300	300	300	300	84.79	4.23
	2	2	2	2	2		
a9a_100_s2	4	4	4	4	4	84.75	0.33
astro-ph_62369.dat_100_s1	229	233	241	237	229	95.49	7.94
astro-ph_62369.dat_100_s2	5	5	6	5	5	95.49	5.58
australian_100_s1	300	300	300	300	300	79.42	0.07
	3	5	5	2	5		
australian_100_s2	4	5	4	5	4	79.57	0.01
australian_scale_100_s1	9	9	9	9	9	86.09	0.00
australian_scale_100_s2	4	5	5	4	4	86.09	0.01
breast-cancer_100_s1	5	3	3	5	4	65.01	0.00
breast-cancer_100_s2	2	2	2	2	1	65.01	0.00
breast-cancer_scale_100_s1	300	300	300	300	300	96.78	0.02
	2	1	2	1	2		
breast-cancer_scale_100_s2	4	5	4	6	4	96.78	0.00
cod-rna_100_s1	300	300	300	300	300	88.40	6.86
	1	1	2	2	3		
cod-rna_100_s2	3	3	3	3	3	88.40	0.36
colon-cancer_100_s1	8	9	7	8	9	75.81	0.05
colon-cancer_100_s2	4	4	4	4	5	75.81	0.08
covtype.libsvm.binary.scale_100_s1	46	45	45	44	44	75.68	33.78
covtype.libsvm.binary.scale_100_s2	9	10	10	9	9	75.69	34.93
covtype.libsvm.binary_100_s1	300	300	300	300	300	64.92	223.71
	4	3	4	2	3		
covtype.libsvm.binary_100_s2	5	8	6	7	6	64.95	21.48
diabetes_100_s1	300	300	300	300	300	67.32	0.05
	2	2	1	4	3		
diabetes_100_s2	1	3	5	2	4	67.45	0.00
diabetes_scale_100_s1	300	300	300	300	300	77.21	0.05
	1	1	1	1	1		
diabetes_scale_100_s2	3	3	3	3	3	77.34	0.00
duke_100_s1	11	13	13	13	15	88.64	0.15
duke_100_s2	9	6	7	8	4	90.91	0.23
fourclass_100_s1	75	70	72	74	71	73.32	0.01
fourclass_100_s2	2	2	2	2	2	73.32	0.00
fourclass_scale_100_s1	9	7	9	8	8	68.56	0.00
fourclass_scale_100_s2	2	2	2	2	2	68.45	0.00
german.numer_100_s1	300	300	300	300	300	76.70	0.15
	5	4	5	4	4		
german.numer_100_s2	3	3	4	3	3	76.80	0.01
german.numer_scale_100_s1	300	300	300	300	300	76.80	0.14
	4	4	4	4	4		
german.numer_scale_100_s2	3	4	3	3	4	76.90	0.01

Table 10 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
gisette_scale_100_s1	62	66	62	64	66	96.97	16.27
gisette_scale_100_s2	5	7	5	5	6	97.07	19.19
heart_100_s1	300	300	300	300	300	83.33	0.03
	5	5	4	5	3		
heart_100_s2	4	5	5	4	6	83.33	0.00
heart_scale_100_s1	8	8	8	9	9	82.96	0.00
heart_scale_100_s2	3	3	3	4	3	82.96	0.00
ijcnn1_100_s1	300	300	300	300	300	92.29	6.15
	3	2	3	3	3		
ijcnn1_100_s2	5	5	6	6	6	92.25	0.60
ionosphere_scale_100_s1	300	300	300	300	300	84.05	0.04
	7	8	6	4	5		
ionosphere_scale_100_s2	6	6	5	6	6	83.76	0.01
kdda_100_s1	300	300	300	300	300	86.52	4630.91
	0	0	0	0	0		
kdda_100_s2	11	11	13	13	14	86.49	2564.00
kddb_100_s1	300	300	300	300	300	87.49	12555.66
	1	1	2	1	2		
kddb_100_s2	17	18	18	18	18	87.52	9795.87
leisure.scale_100_s1	300	300	300	300	300	85.04	1051.34
	3	2	1	3	3		
leisure.scale_100_s2	6	7	6	7	9	85.96	358.16
leu_100_s1	7	8	9	9	8	94.74	0.12
leu_100_s2	4	4	5	4	6	97.37	0.15
liver-disorders_100_s1	300	300	300	300	300	68.28	0.01
	1	1	1	1	1		
liver-disorders_100_s2	3	1	2	1	1	68.97	0.00
liver-disorders_scale_100_s1	99	138	153	104	104	73.79	0.00
liver-disorders_scale_100_s2	3	3	2	3	3	74.48	0.00
madelon_100_s1	300	300	300	300	300	57.15	4.11
	3	3	2	3	2		
madelon_100_s2	7	4	4	6	5	57.20	0.79
mushrooms_100_s1	43	72	39	44	59	100.00	0.10
mushrooms_100_s2	5	5	5	5	6	99.98	0.09
news20.binary_100_s1	300	300	300	300	300	96.74	52.06
	2	1	1	1	1		
news20.binary_100_s2	5	5	6	5	4	97.25	11.32
rcv1_test.binary_100_s1	300	300	300	300	300	97.00	170.94
	0	0	0	0	0		
rcv1_test.binary_100_s2	7	10	9	6	7	97.29	74.93
rcv1_train.binary_100_s1	300	300	300	300	300	96.33	3.73
	0	0	0	0	0		
rcv1_train.binary_100_s2	4	6	5	5	5	96.06	1.52
real-sim_100_s1	300	300	300	300	300	96.66	6.33
	0	0	0	0	0		
real-sim_100_s2	7	6	6	6	6	96.74	4.28
skin_nonskin_100_s1	300	289	274	285	300	90.84	35.45

Table 10 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
	0	0	0	0	0		
skin_nonskin_100_s2	2	2	2	2	2	90.84	1.47
sonar_scale_100_s1	300	300	300	300	300	73.56	0.04
	3	4	2	3	4		
sonar_scale_100_s2	7	8	10	8	9	73.08	0.01
splice_100_s1	300	300	300	300	300	80.10	0.24
	2	2	2	2	2		
splice_100_s2	4	4	4	4	4	80.10	0.03
splice_scale_100_s1	36	33	34	33	35	71.00	0.05
splice_scale_100_s2	3	3	3	3	3	71.00	0.03
svmguide1_100_s1	95	100	94	94	93	83.49	0.05
svmguide1_100_s2	5	5	5	5	5	83.46	0.01
svmguide3_100_s1	300	300	300	300	300	79.73	0.20
	6	5	5	6	4		
svmguide3_100_s2	7	6	5	6	4	79.32	0.03
train308.scale_100_s1	300	300	300	300	300	91.08	95.39
	1	0	0	0	1		
train308.scale_100_s2	11	10	10	10	11	91.56	49.36
url_combined_100_s1	180	180	180	181	168	99.44	516.10
url_combined_100_s2	9	7	6	7	5	99.21	402.35
w8a_100_s1	300	300	300	300	300	98.22	2.49
	7	6	5	5	7		
w8a_100_s2	12	11	9	9	9	98.23	0.56
webspam_wc_normalized_unigram.svm_100_s1	300	300	300	300	300	92.71	274.88
	3	1	1	0	1		
webspam_wc_normalized_unigram.svm_100_s2	5	5	4	5	5	92.66	43.16

Table 11: Multi-core LIBLINEAR for logistic regression (`-s 0`: primal Newton). $C = C_{\text{best}}$. Multi-core dual CD is not available for logistic regression.

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
HIGGS_s0	7	7	8	7	7	64.07	451.05
a9a_s0	5	5	5	5	5	84.84	0.23
astro-ph_62369.dat_s0	6	6	5	6	6	96.79	4.38
australian_s0	5	5	5	5	5	67.83	0.01
australian_scale_s0	2	2	2	2	2	87.10	0.01
breast-cancer_s0	1	1	1	1	0	59.44	0.01
breast-cancer_scale_s0	5	5	5	5	5	96.63	0.01
cod-rna_s0	4	4	4	4	4	87.57	0.45
colon-cancer_s0	1	2	1	1	1	80.65	0.05
covtype.libsvm.binary.scale_s0	10	8	8	6	8	75.66	11.16
covtype.libsvm.binary_s0	4	5	3	4	4	61.60	6.46
diabetes_s0	2	2	2	2	2	67.84	0.01
diabetes_scale_s0	3	3	3	3	3	77.86	0.01
duke_s0	3	3	3	3	3	81.82	0.15
fourclass_s0	3	3	3	3	3	72.74	0.01
fourclass_scale_s0	2	2	2	2	2	68.79	0.01
german.numer_s0	3	3	3	3	3	76.30	0.02
german.numer_scale_s0	4	4	4	4	4	77.40	0.02
gisette_scale_s0	6	6	6	6	6	97.17	11.94
heart_s0	4	4	3	4	4	84.44	0.01
heart_scale_s0	2	2	2	2	2	82.96	0.01
ijcnn1_s0	6	6	6	6	6	92.44	0.56
ionosphere_scale_s0	5	6	5	5	5	83.48	0.02
kdda_s0	7	9	7	7	8	88.23	969.02
kddb_s0	11	14	8	9	12	88.89	2961.35
leisure.scale_s0	8	15	8	10	8	87.15	234.50
leu_s0	6	6	6	6	6	92.11	0.16
liver-disorders_s0	3	3	3	2	2	68.28	0.01
liver-disorders_scale_s0	2	3	2	2	2	73.79	0.01
madelon_s0	1	2	1	1	2	59.60	0.29
mushrooms_s0	6	6	6	6	6	99.96	0.09
news20.binary_s0	6	7	8	7	6	96.91	12.80
rcv1_test.binary_s0	7	7	7	7	7	97.74	43.95
rcv1_train.binary_s0	6	6	6	6	6	96.91	1.40
real-sim_s0	6	6	6	6	7	97.46	3.18
skin_nonskin_s0	4	4	4	4	4	90.67	1.35
sonar_scale_s0	5	6	5	4	4	69.71	0.02
splice_s0	3	3	3	3	3	79.10	0.03
splice_scale_s0	2	2	2	2	2	71.80	0.03
svmguide1_s0	5	5	5	5	5	83.20	0.02
svmguide3_s0	7	5	6	5	5	80.13	0.03
train308.scale_s0	7	7	7	7	7	92.70	24.50
url_combined_s0	5	6	5	5	7	98.09	201.33
w8a_s0	11	9	9	9	10	98.38	0.45
webspam_wc_normalized_unigram.svm_s0	5	5	5	5	5	92.33	20.14

Table 12: Multi-core LIBLINEAR for logistic regression (`-s 0`: primal Newton). $C = 100C_{\text{best}}$. Multi-core dual CD is not available for logistic regression.

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
HIGGS_100_s0	6	7	7	7	7	64.09	426.36
a9a_100_s0	6	5	6	5	5	84.83	0.23
astro-ph_62369.dat_100_s0	7	7	7	7	7	96.25	4.91
australian_100_s0	5	5	5	5	5	78.41	0.02
australian_scale_100_s0	4	4	5	5	4	87.10	0.01
breast-cancer_100_s0	1	1	1	1	1	65.01	0.01
breast-cancer_scale_100_s0	5	5	5	5	5	96.63	0.01
cod-rna_100_s0	4	4	4	4	4	89.21	0.41
colon-cancer_100_s0	4	4	4	4	4	72.58	0.07
covtype.libsvm.binary.scale_100_s0	9	6	8	10	7	75.61	11.26
covtype.libsvm.binary_100_s0	8	13	10	6	9	71.24	12.90
diabetes_100_s0	3	3	2	3	3	69.79	0.01
diabetes_scale_100_s0	4	5	4	3	4	77.73	0.01
duke_100_s0	5	5	5	5	5	88.64	0.17
fourclass_100_s0	3	3	3	3	3	73.32	0.01
fourclass_scale_100_s0	3	3	3	3	3	68.68	0.01
german.numer_100_s0	3	4	4	4	4	76.10	0.02
german.numer_scale_100_s0	4	4	4	4	4	77.50	0.02
gisette_scale_100_s0	7	6	6	6	8	97.17	12.42
heart_100_s0	5	5	7	4	5	84.44	0.01
heart_scale_100_s0	4	4	4	4	4	82.96	0.01
ijcnn1_100_s0	6	7	6	6	6	92.43	0.60
ionosphere_scale_100_s0	5	7	5	5	5	83.48	0.02
kdda_100_s0	7	11	11	7	7	86.65	1078.10
kddb_100_s0	7	7	7	10	8	87.80	1527.79
leisure.scale_100_s0	14	9	11	11	9	86.22	337.42
leu_100_s0	6	6	6	6	6	92.11	0.16
liver-disorders_100_s0	3	3	3	1	3	67.59	0.01
liver-disorders_scale_100_s0	3	4	3	4	4	72.41	0.01
madelon_100_s0	6	6	7	5	5	58.55	0.43
mushrooms_100_s0	6	6	6	6	6	99.98	0.08
news20.binary_100_s0	15	11	11	16	8	97.31	27.43
rcv1_test.binary_100_s0	7	7	7	6	6	97.59	42.63
rcv1_train.binary_100_s0	6	6	6	6	6	96.60	1.26
real-sim_100_s0	7	7	7	7	7	97.42	3.67
skin_nonskin_100_s0	4	4	4	4	4	90.68	1.43
sonar_scale_100_s0	10	11	10	9	8	74.52	0.03
splice_100_s0	4	4	4	4	4	79.90	0.03
splice_scale_100_s0	3	3	3	3	3	69.90	0.03
svmguide1_100_s0	5	5	5	5	5	83.39	0.02
svmguide3_100_s0	4	5	7	5	6	80.45	0.04
train308.scale_100_s0	8	8	8	8	8	92.07	32.40
url_combined_100_s0	6	7	7	6	6	99.02	240.52
w8a_100_s0	15	10	9	11	10	98.37	0.50
webspam_wc_normalized_unigram.svm_100_s0	7	7	6	7	7	92.74	23.52

Table 13: Multi-core LIBLINEAR for L2-loss SVM. (`-s 1`: dual CD, `-s 2`: primal Newton). $C = C_{\text{best}}$. For dual CD, if # iterations exceeds 300, primal Newton is called and the # of Newton iterations is shown in the next row.

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
HIGGS_s1	7	7	7	8	6	64.01	272.06
HIGGS_s2	7	8	7	8	9	63.98	430.98
a9a_s1	15	14	15	15	15	84.84	0.41
a9a_s2	4	4	4	4	4	84.78	0.19
astro-ph_62369.dat_s1	8	8	8	8	8	96.89	2.83
astro-ph_62369.dat_s2	5	5	5	5	5	96.90	3.24
australian_s1	300	300	300	300	300	67.68	0.07
	2	1	0	1	2		
australian_s2	5	5	4	5	5	67.54	0.01
australian_scale_s1	4	4	5	5	4	86.67	0.01
australian_scale_s2	2	2	2	2	2	86.81	0.01
breast-cancer_s1	1	1	1	1	1	59.44	0.00
breast-cancer_s2	1	1	1	1	1	59.44	0.01
breast-cancer_scale_s1	24	26	28	28	25	96.78	0.01
breast-cancer_scale_s2	4	4	4	4	4	96.78	0.01
cod-rna_s1	55	49	54	54	52	87.55	1.31
cod-rna_s2	3	3	3	3	3	87.57	0.30
colon-cancer_s1	3	3	3	3	3	80.65	0.05
colon-cancer_s2	1	2	1	1	1	80.65	0.05
covtype.libsvm.binary.scale_s1	6	6	6	6	6	75.69	6.82
covtype.libsvm.binary.scale_s2	6	8	7	7	7	75.70	8.86
covtype.libsvm.binary_s1	25	25	24	25	25	61.19	16.50
covtype.libsvm.binary_s2	3	2	2	3	3	61.21	4.49
diabetes_s1	300	300	138	274	209	67.71	0.06
	1	1	0	0	0		
diabetes_s2	2	2	2	3	2	67.71	0.01
diabetes_scale_s1	10	10	9	10	9	77.34	0.01
diabetes_scale_s2	3	3	3	3	3	77.34	0.01
duke_s1	6	5	8	8	8	84.09	0.14
duke_s2	4	2	4	3	3	84.09	0.15
fourclass_s1	6	6	6	7	6	72.74	0.01
fourclass_s2	2	2	2	2	2	72.62	0.01
fourclass_scale_s1	4	5	4	4	4	68.56	0.00
fourclass_scale_s2	1	1	1	1	1	68.56	0.00
german.numer_s1	300	300	300	300	300	75.90	0.30
	2	2	2	2	2		
german.numer_s2	3	3	3	3	3	75.90	0.02
german.numer_scale_s1	112	114	82	105	105	77.80	0.09
german.numer_scale_s2	3	4	3	4	4	77.90	0.02
gisette_scale_s1	26	26	26	25	25	97.13	12.49
gisette_scale_s2	6	6	7	5	5	97.18	9.73
heart_s1	300	300	300	300	300	84.81	0.04
	3	4	4	3	4		
heart_s2	2	3	3	3	3	84.81	0.01

Table 13 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
heart_scale_s1	4	4	4	4	3	83.33	0.01
heart_scale_s2	2	2	2	2	2	82.96	0.01
ijcnn1_s1	16	18	17	17	17	92.25	0.69
ijcnn1_s2	5	5	6	6	6	92.26	0.49
ionosphere_scale_s1	300	300	300	300	300	84.05	0.07
	0	1	2	1	1		
ionosphere_scale_s2	4	5	5	4	5	84.05	0.02
kdda_s1	10	10	10	10	10	88.25	376.98
kdda_s2	13	14	14	16	13	88.24	1108.22
kddb_s1	25	25	25	24	25	88.95	1013.04
kddb_s2	15	16	16	14	16	88.95	2207.53
leisure.scale_s1	40	39	39	41	39	87.27	256.36
leisure.scale_s2	8	8	7	9	9	87.27	160.86
leu_s1	8	6	8	8	11	92.11	0.14
leu_s2	6	5	5	5	6	92.11	0.16
liver-disorders_s1	11	11	12	13	12	69.66	0.00
liver-disorders_s2	2	2	2	1	1	69.66	0.01
liver-disorders_scale_s1	7	6	6	7	7	73.79	0.01
liver-disorders_scale_s2	2	2	2	2	2	73.79	0.01
madelon_s1	9	9	9	9	9	59.65	0.39
madelon_s2	1	1	1	1	1	59.60	0.25
mushrooms_s1	38	37	58	41	33	100.00	0.11
mushrooms_s2	5	4	5	5	4	99.95	0.07
news20.binary_s1	300	300	300	300	300	96.84	36.58
	0	0	0	1	0		
news20.binary_s2	7	6	6	5	6	97.24	9.33
rcv1_test.binary_s1	10	10	10	10	10	97.80	36.82
rcv1_test.binary_s2	6	6	6	5	5	97.79	32.15
rcv1_train.binary_s1	10	12	10	10	11	97.05	1.13
rcv1_train.binary_s2	6	6	5	5	6	97.03	1.05
real-sim_s1	9	9	9	9	8	97.51	2.45
real-sim_s2	5	6	6	5	6	97.54	2.41
skin_nonskin_s1	8	8	8	8	8	90.90	1.57
skin_nonskin_s2	2	2	2	2	2	90.90	0.93
sonar_scale_s1	16	18	17	16	17	71.63	0.01
sonar_scale_s2	4	5	4	4	4	72.12	0.01
splice_s1	14	14	13	14	14	79.50	0.04
splice_s2	3	3	3	3	3	79.60	0.02
splice_scale_s1	5	5	5	5	5	71.60	0.03
splice_scale_s2	2	2	2	2	2	71.60	0.02
svmguide1_s1	6	6	6	6	6	83.17	0.02
svmguide1_s2	4	4	5	4	5	83.20	0.02
svmguide3_s1	300	300	300	300	300	79.73	0.35
	3	3	4	3	2		
svmguide3_s2	4	4	5	5	5	79.57	0.03
train308.scale_s1	10	9	9	10	9	92.82	14.66
train308.scale_s2	9	9	8	8	11	92.81	22.01

Table 13 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
url_combined_s1	7	8	7	7	7	98.53	108.90
url_combined_s2	6	6	6	7	7	98.41	159.21
w8a_s1	300	300	300	300	300	98.26	2.75
	2	2	2	2	2		
w8a_s2	8	8	9	8	8	98.26	0.28
webspam_wc_normalized_unigram.svm_s1	11	11	11	11	11	92.61	18.57
webspam_wc_normalized_unigram.svm_s2	4	4	4	4	4	92.54	17.38

Table 14: Multi-core LIBLINEAR for L2-loss SVM. (`-s 1`: dual CD, `-s 2`: primal Newton). $C = 100C_{\text{best}}$. For dual CD, if # iterations exceeds 300, primal Newton is called and the # of Newton iterations is shown in the next row.

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
HIGGS_100_s1	300	300	300	300	175	64.04	2141.30
	1	1	1	1	0		
HIGGS_100_s2	7	8	8	9	9	64.03	468.85
a9a_100_s1	300	300	300	300	300	84.76	5.77
	2	2	2	2	2		
a9a_100_s2	4	4	4	4	4	84.81	0.19
astro-ph_62369.dat_100_s1	240	237	115	215	241	95.49	6.48
astro-ph_62369.dat_100_s2	5	5	5	5	5	95.48	3.22
australian_100_s1	300	300	300	300	300	78.70	0.12
	4	4	3	2	4		
australian_100_s2	4	5	4	5	4	79.13	0.01
australian_scale_100_s1	10	9	9	9	10	85.80	0.01
australian_scale_100_s2	4	5	5	4	3	85.80	0.01
breast-cancer_100_s1	1	1	1	1	1	59.44	0.00
breast-cancer_100_s2	1	2	2	2	2	65.01	0.01
breast-cancer_scale_100_s1	300	300	300	300	300	96.78	0.04
	3	1	2	3	2		
breast-cancer_scale_100_s2	4	4	4	4	4	96.78	0.01
cod-rna_100_s1	300	300	300	300	300	88.33	12.76
	1	2	1	2	2		
cod-rna_100_s2	3	3	3	3	3	88.33	0.30
colon-cancer_100_s1	10	10	11	10	10	69.35	0.06
colon-cancer_100_s2	3	4	4	4	4	69.35	0.06
covtype.libsvm.binary.scale_100_s1	49	49	49	49	50	75.65	44.80
covtype.libsvm.binary.scale_100_s2	10	12	11	11	9	75.67	12.51
covtype.libsvm.binary_100_s1	300	300	300	300	300	64.89	270.54
	4	4	3	3	3		
covtype.libsvm.binary_100_s2	6	5	6	5	7	64.89	7.00
diabetes_100_s1	300	300	300	300	300	69.40	0.11
	3	3	3	4	4		
diabetes_100_s2	6	1	1	5	2	69.27	0.01
diabetes_scale_100_s1	300	300	300	300	300	77.86	0.10
	1	1	1	1	1		
diabetes_scale_100_s2	3	3	3	3	3	77.60	0.01
duke_100_s1	13	15	18	13	13	81.82	0.18
duke_100_s2	6	5	5	6	10	88.64	0.18
fourclass_100_s1	117	72	119	76	117	73.09	0.03
fourclass_100_s2	2	2	2	2	2	73.20	0.01
fourclass_scale_100_s1	9	9	9	10	9	68.68	0.01
fourclass_scale_100_s2	2	2	1	2	2	68.68	0.01
german.numer_100_s1	300	300	300	300	300	76.40	0.28
	4	4	3	4	4		
german.numer_100_s2	3	3	4	4	3	76.60	0.02
german.numer_scale_100_s1	300	300	300	300	300	77.70	0.26
	3	4	4	4	4		

Table 14 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
german.numer_scale_100_s2	3	3	3	4	3	77.90	0.02
gisette_scale_100_s1	64	67	68	69	68	96.85	16.17
gisette_scale_100_s2	5	5	7	8	6	97.08	9.42
heart_100_s1	300	300	300	300	300	85.19	0.04
	3	4	4	4	6		
heart_100_s2	5	4	5	4	5	85.19	0.01
heart_scale_100_s1	9	9	9	10	9	82.96	0.01
heart_scale_100_s2	3	3	4	3	3	82.96	0.01
ijcnn1_100_s1	300	300	300	300	300	92.27	10.53
	3	3	3	2	3		
ijcnn1_100_s2	6	6	6	6	7	92.27	0.49
ionosphere_scale_100_s1	300	300	300	300	300	83.76	0.07
	7	5	6	7	5		
ionosphere_scale_100_s2	5	5	5	5	5	84.33	0.02
kdda_100_s1	300	300	300	300	300	86.52	6541.42
	1	0	1	1	1		
kdda_100_s2	11	13	11	11	14	86.47	991.36
kddb_100_s1	300	300	300	300	300	87.49	10319.54
	3	2	2	2	2		
kddb_100_s2	17	12	12	18	18	87.59	2509.10
leisure.scale_100_s1	300	300	300	300	300	84.93	1526.87
	2	3	2	2	3		
leisure.scale_100_s2	7	8	8	7	7	85.74	138.20
leu_100_s1	9	6	7	8	9	92.11	0.13
leu_100_s2	4	3	5	3	7	92.11	0.13
liver-disorders_100_s1	300	300	300	300	300	67.59	0.03
	2	1	2	2	2		
liver-disorders_100_s2	2	3	1	2	1	67.59	0.01
liver-disorders_scale_100_s1	146	148	136	208	138	71.72	0.01
liver-disorders_scale_100_s2	2	4	3	3	3	71.72	0.01
madelon_100_s1	300	300	300	300	300	58.45	4.99
	3	3	3	2	3		
madelon_100_s2	4	4	6	5	4	58.35	0.36
mushrooms_100_s1	43	45	98	51	46	100.00	0.11
mushrooms_100_s2	5	5	5	5	5	99.96	0.06
news20.binary_100_s1	300	300	300	300	300	96.81	67.25
	2	1	1	1	2		
news20.binary_100_s2	5	5	5	5	5	97.39	7.86
rcv1_test.binary_100_s1	300	300	300	300	300	96.97	266.07
	0	0	0	0	0		
rcv1_test.binary_100_s2	6	6	6	6	6	97.49	31.56
rcv1_train.binary_100_s1	300	300	300	300	300	96.21	8.27
	0	0	0	0	0		
rcv1_train.binary_100_s2	4	4	5	5	5	96.25	0.97
real-sim_100_s1	300	300	300	300	300	96.48	9.34
	0	0	0	0	0		
real-sim_100_s2	6	6	6	6	6	96.62	2.21
skin_nonskin_100_s1	294	277	279	269	300	90.90	40.29

Table 14 – continued from previous page

Data set and approaches	# Iter	# Iter	# Iter	# Iter	# Iter	CV	Time
	0	0	0	0	1		
skin_nonskin_100_s2	2	2	2	2	2	90.90	0.88
sonar_scale_100_s1	300	300	300	300	300	73.08	0.06
	3	7	3	3	3		
sonar_scale_100_s2	10	11	9	11	10	72.60	0.03
splice_100_s1	300	300	300	300	300	80.20	0.39
	2	2	3	3	2		
splice_100_s2	4	4	4	4	4	80.10	0.03
splice_scale_100_s1	35	34	34	36	35	70.60	0.07
splice_scale_100_s2	3	4	3	3	3	70.80	0.03
svmguide1_100_s1	100	107	102	98	102	83.43	0.11
svmguide1_100_s2	4	4	5	5	5	83.36	0.02
svmguide3_100_s1	300	300	300	300	300	79.73	0.32
	7	7	6	5	3		
svmguide3_100_s2	5	5	6	5	8	79.81	0.03
train308.scale_100_s1	300	300	300	300	300	91.10	111.62
	2	1	1	1	2		
train308.scale_100_s2	12	12	12	13	9	91.57	24.51
url_combined_100_s1	178	181	177	180	164	99.43	285.93
url_combined_100_s2	7	6	7	7	6	99.27	159.96
w8a_100_s1	300	300	300	300	300	98.26	2.81
	5	7	6	5	7		
w8a_100_s2	10	9	9	10	8	98.27	0.29
webspam_wc_normalized_unigram.svm_100_s1	300	300	300	300	300	92.72	228.85
	3	1	1	1	1		
webspam_wc_normalized_unigram.svm_100_s2	5	5	5	5	7	92.69	18.71