The 12th International Conference on Intelligent Data Processing: Theory and Applications

IDP-2018

October, 8 – 12th, 2018, Gaeta, Italy

Fast approximate two-class SVM learning in the case of a large number of objects

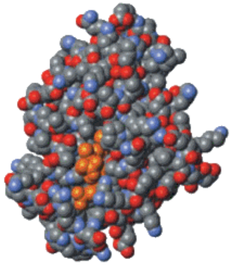
***Makarova Alexandra**, Sulimova Valentina*

aleksarova@gmail.com, vsulimova@yandex.ru,

Russia, Tula State University
Laboratory of Data Analysis

Two-class recognition problem

Mass sources of applied problems:
- molecular biology;
- medical systems;
- video surveillance systems;
- marketing;
- text analysis;
- biometric verification of personality;
- etc.



An important feature of modern applied recognition problems:
- large amounts of data that require processing

The formulation of the problem of learning two-class recognition

Given:

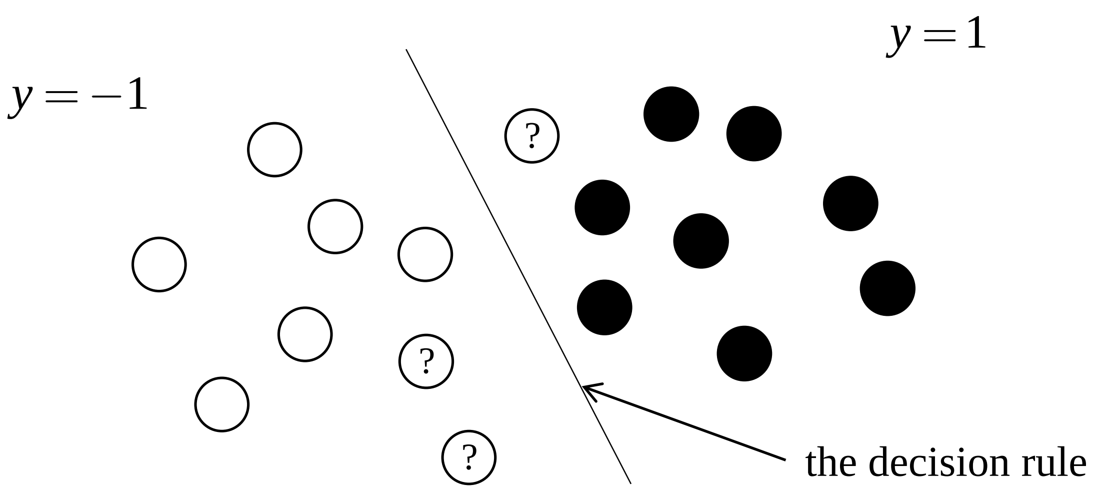
Ω - the set of all possible objects $\omega \in \Omega$ of an arbitrary kind

The hidden function of class affiliation: $y(\omega): \Omega \to \{+1; -1\}$

Training set : $\{(\omega_j, y_j), j = 1, \ldots, N\}, \quad \omega_j \in \Omega^* \subset \Omega, \quad y_j = y(\omega_j) \in \{+1; -1\}$

Required:

to build a decision rule for assigning any objects to one of two classes $\hat{y}(\omega) = \pm 1$



Support Vector Machine (SVM).
Learning in a linear feature space

Representation of objects as points in *m*-dimensional feature space: $\mathbf{x}(\omega) \in R^m$

Training set: $\{\mathbf{x}_j, y_j\}, \quad \mathbf{x}_j = \mathbf{x}_j(\omega), \quad j = 1,...,N$

The decision rule in the form of a linear separating hyperplane:

$$d(\mathbf{x};\mathbf{a},b) = \mathbf{a}^T\mathbf{x} + b \quad \begin{array}{l} \geq 0 \Rightarrow \hat{y}(\mathbf{x}) = +1, \\ < 0 \Rightarrow \hat{y}(\mathbf{x}) = -1, \end{array}$$

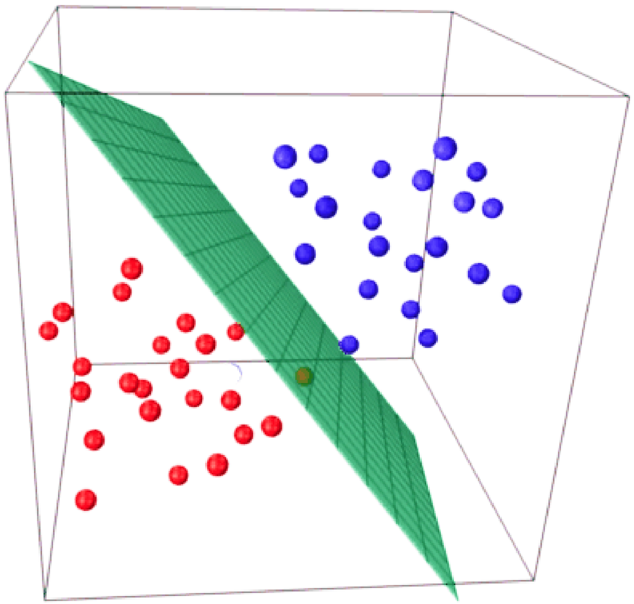
$\mathbf{a} \in R^m$ - direction vector

b - offset along the direction vector

Parameters of the optimal hyperplane
(*in terms of Lagrange multipliers)* $\lambda_j,\, j = 1,...,N:$

$$\mathbf{a} = \sum_{j=1}^{N} \lambda_j y_j \mathbf{x}_j$$

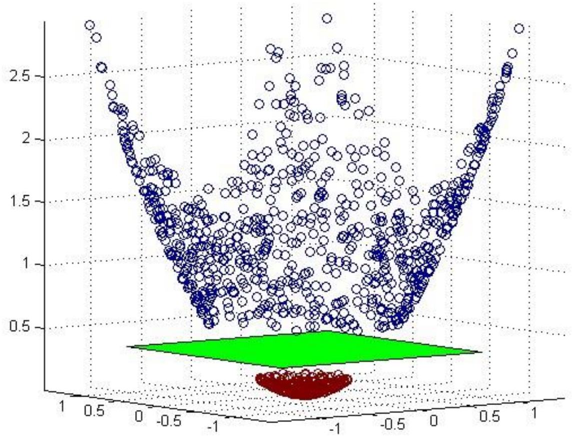
$$b = \frac{1}{2}\left[\min_{j:y_j=1} \sum_{k=1}^{N} \lambda_k y_k \mathbf{x}_j^T \mathbf{x}_k - \max_{j:y_j=-1} \sum_{k=1}^{N} \lambda_k y_k \mathbf{x}_j^T \mathbf{x}_k\right]$$



Support Vector Machine (SVM).
Learning in a space generated by the potential function

Potential function $K(\omega', \omega'')$ - function of object's similarities $\omega', \omega'' \in \Omega$, whose matrix of values is non-negatively defined for any finite set of objects.

$K(\omega', \omega'')$ immerses set of objects Ω into a hypothetical linear space $\tilde{\Omega} \supset \Omega$ in which it plays the role of scalar product

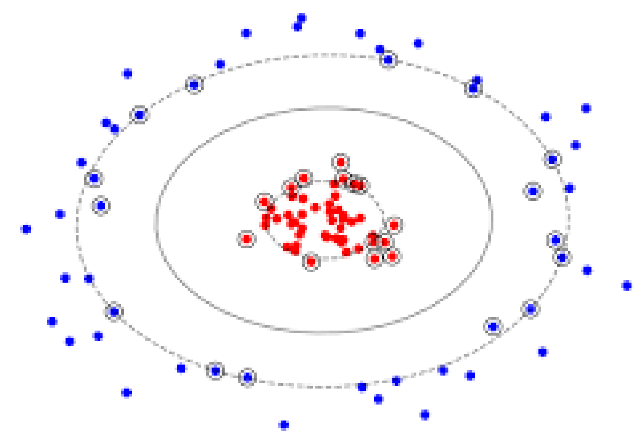


The most popular potential function (radial):

$$K(\omega', \omega'') = \exp[-\gamma \, \| \mathbf{x}(\omega') - \mathbf{x}(\omega'') \|^2]$$

Decision rule: $d(\omega; \mathbf{a}, b) = \sum_{j=1}^{N} a_j K(\omega_j, \omega) + b$

Parameters of the optimal hyperplane
(*in terms of Lagrange multipliers*) $\lambda_j, \; j = 1, \ldots, N:$



$$\mathbf{a} = \sum_{j=1}^{N} \lambda_j y_j \mathbf{x}_j$$

$$b = \frac{1}{2} \left[\min_{j: y_j = 1} \sum_{k=1}^{N} \lambda_k y_k K(\omega_j, \omega_k) - \max_{j: y_j = -1} \sum_{k=1}^{N} \lambda_k y_k K(\omega_j, \omega_k) \right]$$

Many approaches to solving the SVM problem

1. Bottou L. (2004). Stochastic Learning. *Advanced Lectures on Machine Learning*, 146-168, Edited by Olivier Bousquet and Ulrike von Luxburg, Lecture Notes in Artificial Intelligence, LNAI 3176, Springer Verlag, Berlin, 2004.

2. Boser, B. E., Guyon, I. M., Vapnik, V. (1992). A Training Algorithm for Optimal Margin Classifiers, Fifth Annual Workshop on Computational Learning Theory, ACM, 1992

3. Platt, J. (1998). Sequential minimal optimization: A fast algorithm for training support vector machines (Technical Report MSR-TR-98-14). Microsoft Research.

4. Chang, C.-C., & Lin, C.-J. (2001). LIBSVM: a library for support vector machines. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.

5. Joachims T.(1999) “Making large-scale support vector machine learning practical” in Advances in kernel methods: support vector learning. Cambridge, MA, USA: MIT Press, 1999, pp. 169–184

6.Chu, C.-T., Kim, S. K., Lin, Y.-A., Yu, Y., Bradski, G., Ng, A. Y., & Olukotun, K. (2006). Map reduce for machine learning on multicore. NIPS

7. Zhao H. X., Magoules F. (2011). Parallel support vector machines on multi-core and multiprocessor systems. R. Fox. 11th International Conference on Arti_cial Intelligence and Applications (AIA 2011), Feb 2011, Innsbruck, Austria. IASTED, 2011, <10.2316/P.2011.717-056>

8. Graf, H. P., Cosatto, E., Bottou, L., Dourdanovic, I., & Vapnik, V. (2005). Parallel support vector machines: The cascade svm. In Advances in neural information processing systems 17, 521–528.

9. Hsieh C.-J., Dhillon S. Si, and I. S. (2013). A divide-and-conquer solver for kernel support vector machines. arXiv preprint arXiv:1311.0914, 2013.

10. Yang You, James Demmel, Kenneth Czechowski, Le Song and Richard Vuduc (2015). CA-SVM: Communication-Avoiding Parallel Support Vector Machines on Distributed Systems. EECS Department University of California, Berkeley Technical Report No. UCB/EECS-2015-9 February 27, 2015

11. A. Athanasopoulos, A. Dimou, V. Mezaris, I. Kompatsiaris, "GPU Acceleration for Support Vector Machines", Proc. 12th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2011), Delft, The Netherlands, April 2011

12. Carpenter, Austin. (2009). cuSVM: a CUDA implementation of support vector classification and regression.

13. etc.

SVM open source implementations

	LibSVM	liblinear	mpi – libliear	SVMlight	SvmSgd	πSVM	CuSVM	GPU – accelerated LibSVM
Programming language	C	C	C+MPI	C	C	C+MPI	C+CUDA	C+CUDA
Interface	MATLAB, python, R, SciLAB, Java, Perl, Ruby и др.	MATLAB, python, R, SciLAB, Java, Perl, Ruby и др.	–	–	python	–	MATLAB	–
OS	cross – platform	cross – platform	Unix	cross – platform	cross – platform	Unix	Windows	cross – platform
Work in a linear feature space	+	+	+	+	+	+	+	+
Work in a space generated by the potential function	+	–	–	+	–	+	+	+
Work with sparse matrix	–	+	+	+	–	–	–	–
Work on 1 CPU	+	+	+	+	+	+	–	–
Work on several CPU	–	–	+	–	–	+	–	–
Work on GPU	–	–	–	–	–	–	+	+
Software hardware dependency	–	–	+	–	–	+	+	+
Iterative process	+	+	+	+	+	+	+	+

Objective

The study of the basic opportunity of creating an approximate method for solving the SVM problem in the case of a large number of objects, which at the same time is:

- **fast**,
- **economical in memory** (to enable work on a single computer),

and has **a high degree of parallelism** (for the organization of high-performance computing for systems with shared and distributed memory).

Further research will be aimed at developing a version of the proposed algorithm in the space generated by the potential function.

Proposed method of fast approximate solution of the SVM problem

Initial training set: $[X,Y],\ \ X=[\mathbf{x}_j,\ j=1,\dots,N],\quad \mathbf{x}_j\in R^m$

$$Y=[y_j,\ j=1,\dots,N],\quad y_j\in\{-1;1\}$$

Set of random subsamples from the training set: $[X,Y]^{(i)}\in[X,Y],\quad i=1,\dots,k$

The result of training for the i^{th} subsample: $[X,Y]^{(i)}$

$$[\mathbf{a}^{(i)},b^{(i)}],\quad i=1,\dots,k$$

Averaged decision rule: $[\mathbf{a},b],\quad \mathbf{a}=\dfrac{1}{k}\sum_{i=1}^{k}\mathbf{a}^{(i)},\quad b=\dfrac{1}{k}\sum_{i=1}^{k}b^{(i)}$

Justification:

$[\mathbf{a}^{(i)},b^{(i)}],\quad i=1,\dots,k$ - random variables with characteristics m and variance d

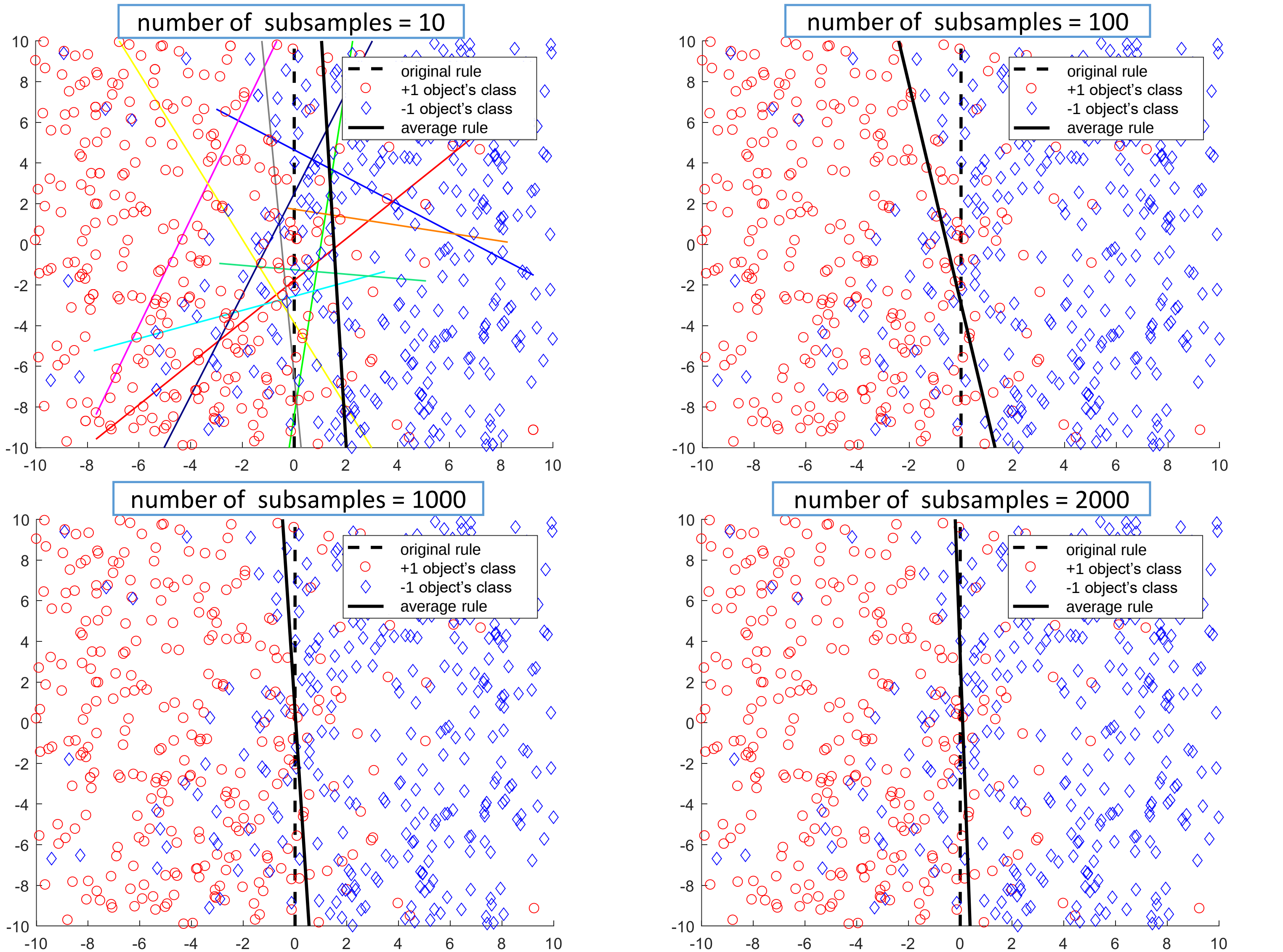
Then $[\mathbf{a},b]$ - random variables with characteristics $M[\mathbf{a},b]=m$ and $D[\mathbf{a},b]=\dfrac{d}{k}$

According to the law of large numbers, the averaged estimate of the parameters of the decision rule converges by probabilities to the mathematical expectation of the corresponding random variable:

$$\lim_{k\to\infty}P\left\{\left|\frac{1}{k}\sum_{i=1}^{k}[\mathbf{a}^{(i)},b^{(i)}]-m\right|<\varepsilon\right\}=1$$

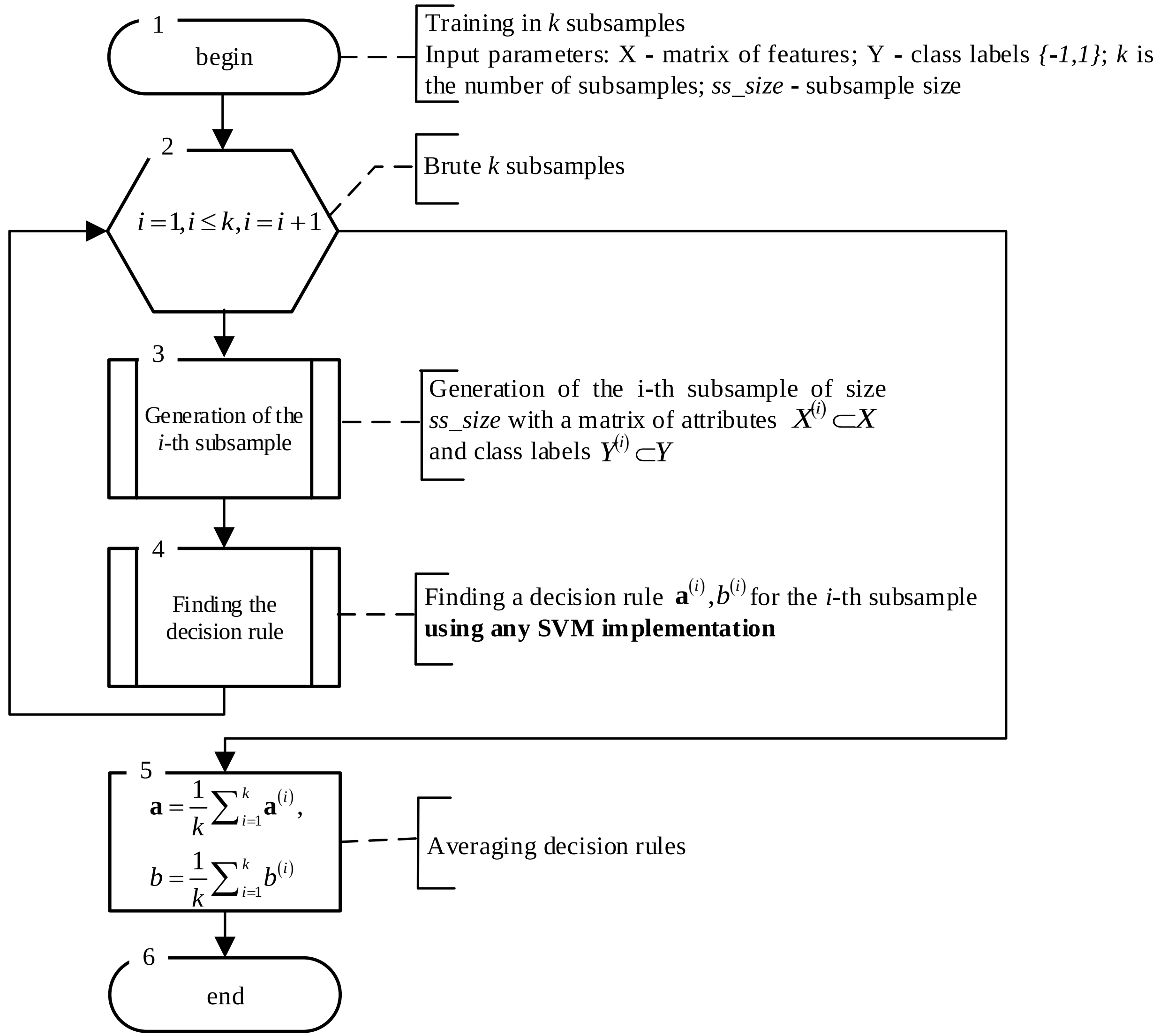
Consequently, with an increasing number of subsamples **the averaged decision rule stabilizes and, in the limit, ceases to be a random variable.**

Position change of the averaged hyperplane in number change of subsamples



display of the sample generated for the two-dimensional case (300 objects of each class, degree of mixing of classes c = 0.8) and decision rules that divided into random intersecting subsamples

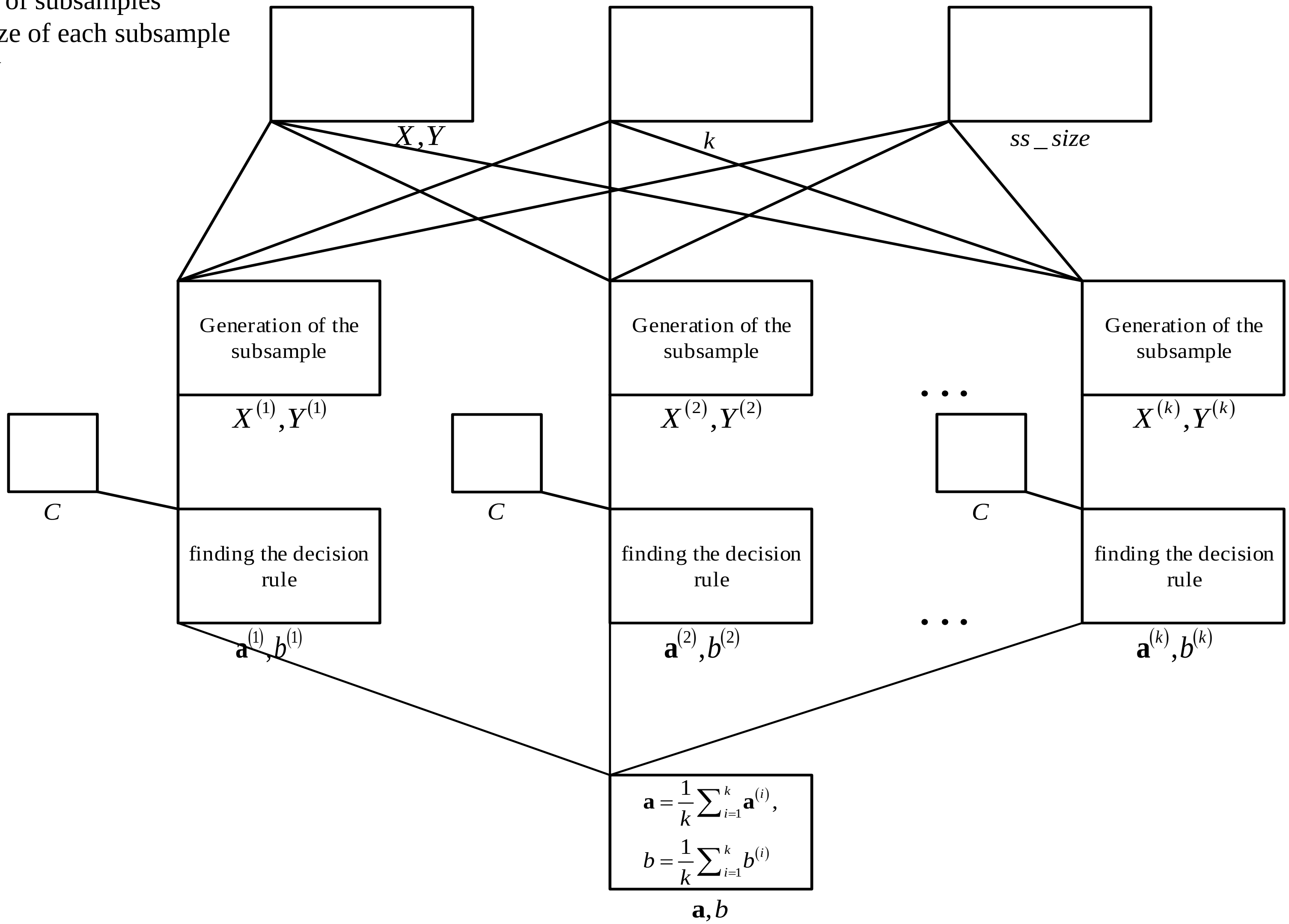
Sequential algorithm
for fast, approximate solution of the SVM problem



Computational scheme of the proposed approach

The model of parallel computing in the form of a graph of operation-operands **for an infinite number of processes** (vertices are operations, edges are data connections).

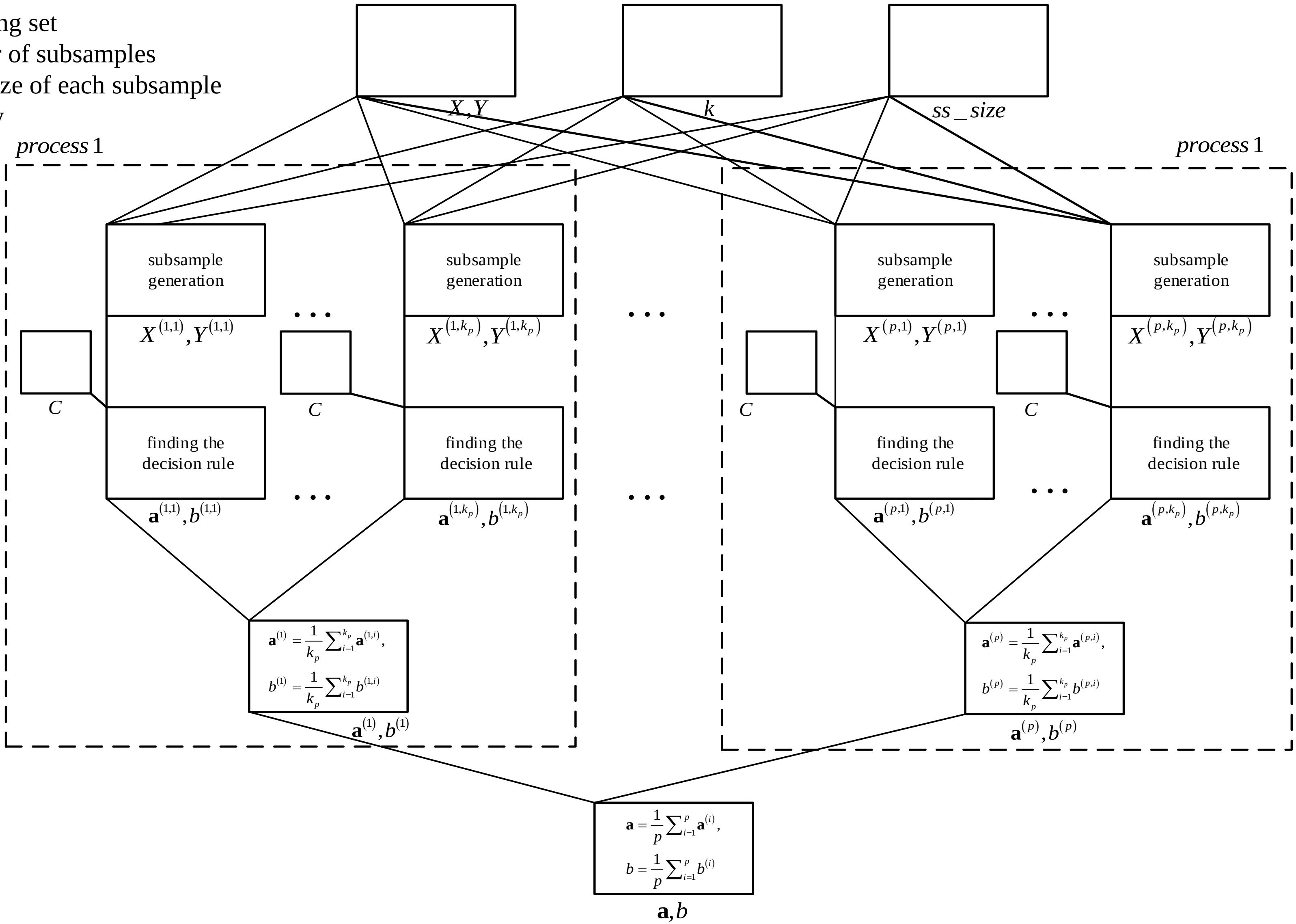
XY – training set
k – number of subsamples
ss_size – size of each subsample
C – penalty



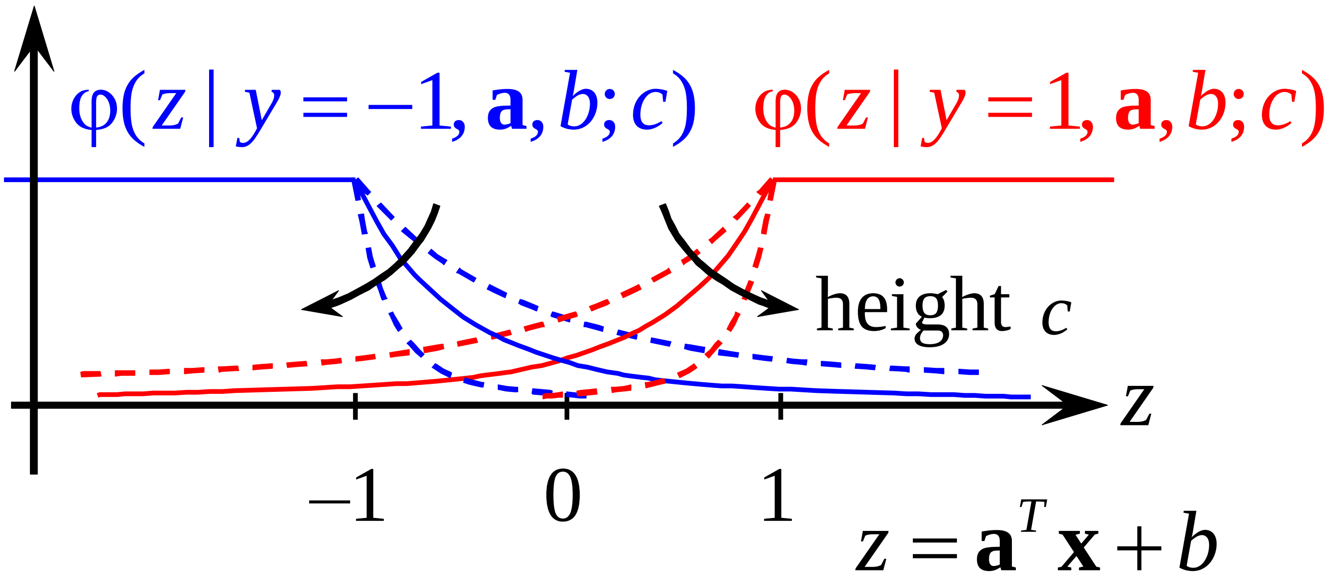
Computational scheme of the proposed approach

The model of parallel computing in the form of a graph of operation-operands **for *p* processes** (vertices are operations, edges are data connections).

XY – training set
k – number of subsamples
ss_size – size of each subsample
C – penalty



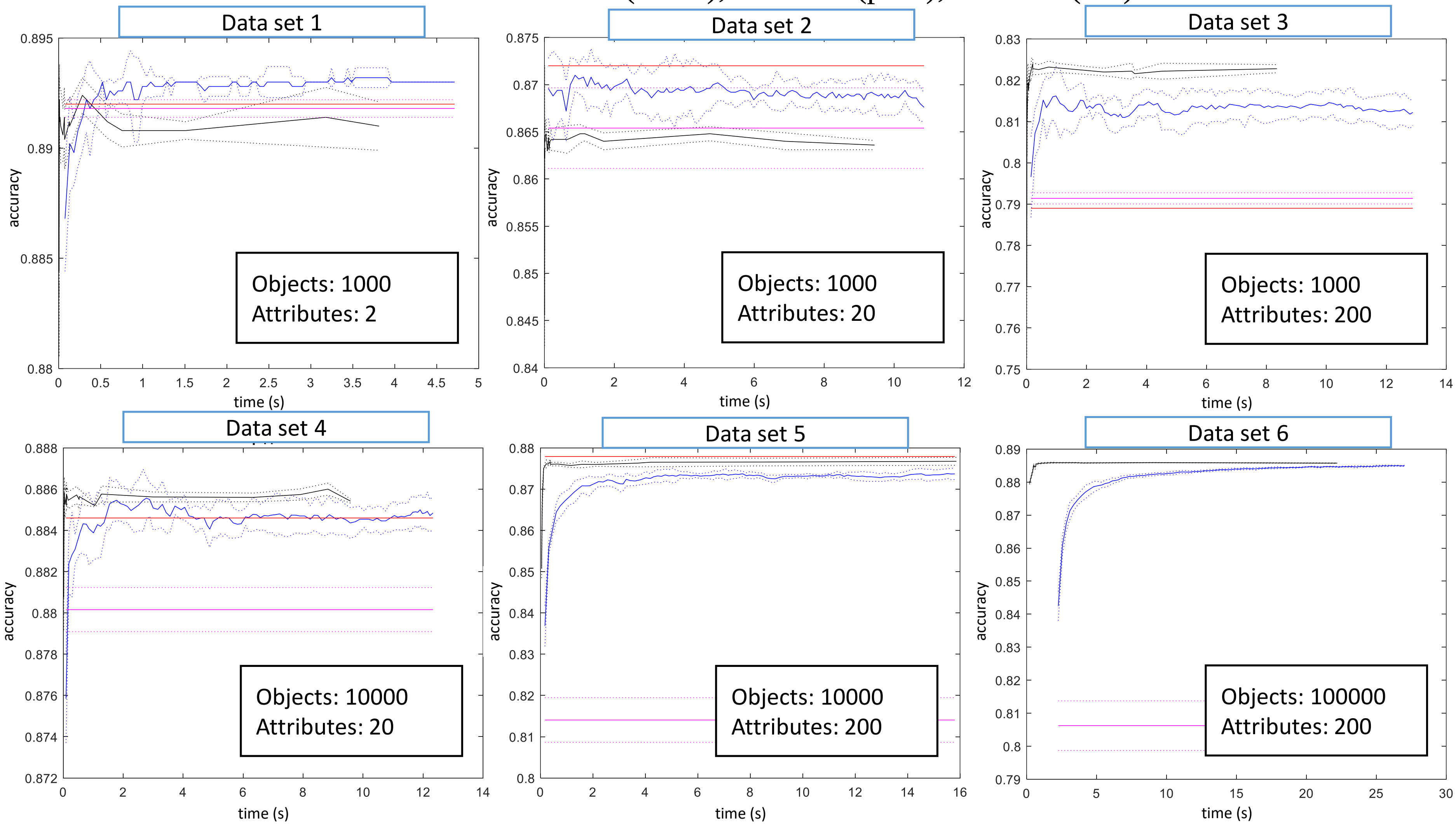
Model data generation in accordance with the probabilistic model of SVM*



* A. Tatarchuk Bayesian support vector machine for learning pattern recognition with controlled selectivity of feature selection. Thesis of Ph.D. n Computing Center RAS, 2014

Experimental results for the model data

Comparison of the proposed method (blue)
with the SGD method (black); liblinear (pink); libSVM (red).



Number of subsamples: 50 - 5000 in increments of 50; subsample size: 50

Results of testing on real data

	SVMlight* (C)		libSVM* (C)		SVC** (LibSVM, python)		πSVM* (1 process)		liblinear* (C)		linear-SVC** (Liblinear, python)		MPI - liblinear* (1 process)		SGD** (python)		proposed method**	
	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy
ijcnn1																		
train	12,41	0,917	22,193	0,917	40,4	0,917	13,99	0,917	0,16	0,913	0,9151	0,9164	0,45	0,913	9,89	0,905	0,412	0,9147
test	0,0267		21,146		25,72		24,67		0,34		0,02		0,09		0,01		0,004	
mnist_576																		
train	51,26	0,994	512,736	0,994	499,8	0,994	95,41	0,994	2,17	0,989	4,5533	0,9935	4,73	0,992	0,51	0,903	5,74	0,987
test	0,0067		30,506		44,08		35,12		1,03		0,0156		1,09		0,02		0,005	
mnist_784																		
train	39,916	0,967	838,573	0,967	717,7	0,967	450,52	0,967	2,49	0,948	5,5304	0,9672	5,37	0,948	0,77	0,903	5,65	0,945
test	0,03		64,767		50,13		68,45		1,27		0,0156		1,32		0,03		0,005	
covtype_2vr																		
train	>40000	-	>180000	-	>3600	-	>40000	-	36,39	0,755	238,44	0.545	0,79	0,612	185	0,505	987	0,6347
test	-		-		-		-		0,69		0,1718		0,8		0,17		0,013	
ijcnn1 (with standardization)																		
train	197,21	0,957	200,8	0,957	122,5	0,917	25,96	0,930	0,14	0,602	5,486	0,916	0,08	0,602	7,44	0,905	2,11	0,9232
test	102,73		82,1		1,13		108,78		0,59		0,018		0,58		0,03		0,003	
mnist_576 (with standardization)																		
train	33072,2	0,903	31946,4	0,992	>3600	-	144,98	0,999	73,09	0,822	32,067	0,9905	36,95	0,822	118	0,988	5,31	0,9879
test	1189,63		16,27		-		52,115		1,57		0,047		1,57		0,03		0,004	
mnist_784 (with standardization)																		
train	31083,3	0,903	19446,2	0,903	>3600	-	483,86	0,995	13,98	0,678	62,699	0,9458	13,26	0,679	83,8	0,947	6,256	0,9637
test	1039,41		730,48		-		82,088		1,32		0,0312		1,3		0,02		0,006	
covtype_2vr (with standardization)																		
train	>40000	-	>180000	-	>3600	-	>40000	-	2,72	0,756	157,99	0,7561	2,47	0,754	130	0,693	11,61	0,7534
test	-		-		-		-		4,24		0,0937		4,29		0,08		0,02	

Data description

Set	Objects on train	Objects on recognition	Number of features
ijcnn1	35000	91701	22
mnist-576-rbf-8vr	60000	10000	576
mnist-784-poly-8vr	60000	10000	784
covtype-2vr	300000	281012	54

Characteristics of CS

*The supercomputer complex of MSU “Lomonosov”: 1 node: Intel Xeon X5570 (2.93 GHz), 8 core, 8Gb RAM, node: 5104

** PC: Intel® Core™ i5-4210U (2.4GHz), 2 core, 6Gb RAM

Results of testing on real data

	SVMlight* (C)		libSVM* (C)		SVC** (LibSVM, python)		πSVM* (1 process)		liblinear* (C)		linear-SVC** (Liblinear, python)		MPI - liblinear* (1 process)		SGD** (python)		proposed method**	
	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy
ijcnn1																		
train	12,41	0,917	22,193	0,917	40,4	0,917	13,99	0,917	0,16	0,913	0,9151	0,9164	0,45	0,913	9,89	0,905	0,412	0,9147
test	0,0267		21,146		25,72		24,67		0,34		0,02		0,09		0,01		0,004	
mnist_576																		
train	51,26	0,994	512,736	0,994	499,8	0,994	95,41	0,994	2,17	0,989	4,5533	0,9935	4,73	0,992	0,51	0,903	5,74	0,987
test	0,0067		30,506		44,08		35,12		1,03		0,0156		1,09		0,02		0,005	
mnist_784																		
train	39,916	0,967	838,573	0,967	717,7	0,967	450,52	0,967	2,49	0,948	5,5304	0,9672	5,37	0,948	0,77	0,903	5,65	0,945
test	0,03		64,767		50,13		68,45		1,27		0,0156		1,32		0,03		0,005	
covtype_2vr																		
train	>40000	-	>180000	-	>3600	-	>40000	-	36,39	0,755	238,44	0.545	0,79	0,612	185	0,505	987	0,6347
test	-		-		-		-		0,69		0,1718		0,8		0,17		0,013	
ijcnn1 (with standardization)																		
train	197,21	0,957	200,8	0,957	122,5	0,917	25,96	0,930	0,14	0,602	5,486	0,916	0,08	0,602	7,44	0,905	2,11	0,9232
test	102,73		82,1		1,13		108,78		0,59		0,018		0,58		0,03		0,003	
mnist_576 (with standardization)																		
train	33072,2	0,903	31946,4	0,992	>3600	-	144,98	0,999	73,09	0,822	32,067	0,9905	36,95	0,822	118	0,988	5,31	0,9879
test	1189,63		16,27		-		52,115		1,57		0,047		1,57		0,03		0,004	
mnist_784 (with standardization)																		
train	31083,3	0,903	19446,2	0,903	>3600	-	483,86	0,995	13,98	0,678	62,699	0,9458	13,26	0,679	83,8	0,947	6,256	0,9637
test	1039,41		730,48		-		82,088		1,32		0,0312		1,3		0,02		0,006	
covtype_2vr (with standardization)																		
train	>40000	-	>180000	-	>3600	-	>40000	-	2,72	0,756	157,99	0,7561	2,47	0,754	130	0,693	11,61	0,7534
test	-		-		-		-		4,24		0,0937		4,29		0,08		0,02	

Data description

Set	Objects on train	Objects on recognition	Number of features
ijcnn1	35000	91701	22
mnist-576-rbf-8vr	60000	10000	576
mnist-784-poly-8vr	60000	10000	784
covtype-2vr	300000	281012	54

Characteristics of CS

*The supercomputer complex of MSU “Lomonosov”:
1 node: Intel Xeon X5570 (2.93 GHz), 8 core,
8Gb RAM, node: 5104
** PC: Intel® Core™ i5-4210U (2.4GHz), 2 core,
6Gb RAM

Results of testing on real data

	SVMlight* (C)		libSVM* (C)		SVC** (LibSVM, python)		πSVM* (1 process)		liblinear* (C)		linear-SVC** (Liblinear, python)		MPI - liblinear* (1 process)		SGD** (python)		proposed method**	
	time (s)	accuracy	time (s)	accuracy	time (s)	accuracy	time (s)	accuracy	time (s)	accuracy	time (s)	accuracy	time (s)	accuracy	time (s)	accuracy	time (s)	accuracy
ijcnn1																		
train	12,41	0,917	22,193	0,917	40,4	0,917	13,99	0,917	0,16	0,913	0,9151	0,9164	0,45	0,913	9,89	0,905	0,412	0,9147
test	0,0267		21,146		25,72		24,67		0,34		0,02		0,09		0,01		0,004	
mnist_576																		
train	51,26	0,994	512,736	0,994	499,8	0,994	95,41	0,994	2,17	0,989	4,5533	0,9935	4,73	0,992	0,51	0,903	5,74	0,987
test	0,0067		30,506		44,08		35,12		1,03		0,0156		1,09		0,02		0,005	
mnist_784																		
train	39,916	0,967	838,573	0,967	717,7	0,967	450,52	0,967	2,49	0,948	5,5304	0,9672	5,37	0,948	0,77	0,903	5,65	0,945
test	0,03		64,767		50,13		68,45		1,27		0,0156		1,32		0,03		0,005	
covtype_2vr																		
train	>40000	-	>180000	-	>3600	-	>40000	-	36,39	0,755	238,44	0.545	0,79	0,612	185	0,505	987	0,6347
test	-		-		-		-		0,69		0,1718		0,8		0,17		0,013	
ijcnn1 (with standardization)																		
train	197,21	0,957	200,8	0,957	122,5	0,917	25,96	0,930	0,14	0,602	5,486	0,916	0,08	0,602	7,44	0,905	2,11	0,9232
test	102,73		82,1		1,13		108,78		0,59		0,018		0,58		0,03		0,003	
mnist_576 (with standardization)																		
train	33072,2	0,903	31946,4	0,992	>3600	-	144,98	0,999	73,09	0,822	32,067	0,9905	36,95	0,822	118	0,988	5,31	0,9879
test	1189,63		16,27		-		52,115		1,57		0,047		1,57		0,03		0,004	
mnist_784 (with standardization)																		
train	31083,3	0,903	19446,2	0,903	>3600	-	483,86	0,995	13,98	0,678	62,699	0,9458	13,26	0,679	83,8	0,947	6,256	0,9637
test	1039,41		730,48		-		82,088		1,32		0,0312		1,3		0,02		0,006	
covtype_2vr (with standardization)																		
train	>40000	-	>180000	-	>3600	-	>40000	-	2,72	0,756	157,99	0,7561	2,47	0,754	130	0,693	11,61	0,7534
test	-		-		-		-		4,24		0,0937		4,29		0,08		0,02	

Data description

Set	Objects on train	Objects on recognition	Number of features
ijcnn1	35000	91701	22
mnist-576-rbf-8vr	60000	10000	576
mnist-784-poly-8vr	60000	10000	784
covtype-2vr	300000	281012	54

Characteristics of CS

*The supercomputer complex of MSU “Lomonosov”: 1 node: Intel Xeon X5570 (2.93 GHz), 8 core, 8Gb RAM, node: 5104

** PC: Intel® Core™ i5-4210U (2.4GHz), 2 core, 6Gb RAM

Results of testing on real data

	SVMlight* (C)		libSVM* (C)		SVC** (LibSVM, python)		πSVM* (1 process)		liblinear* (C)		linear-SVC** (Liblinear, python)		MPI - liblinear* (1 process)		SGD** (python)		proposed method**	
	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy
ijcnn1																		
train	12,41	0,917	22,193	0,917	40,4	0,917	13,99	0,917	0,16	0,913	0,9151	0,9164	0,45	0,913	9,89	0,905	0,412	0,9147
test	0,0267		21,146		25,72		24,67		0,34		0,02		0,09		0,01		0,004	
mnist_576																		
train	51,26	0,994	512,736	0,994	499,8	0,994	95,41	0,994	2,17	0,989	4,5533	0,9935	4,73	0,992	0,51	0,903	5,74	0,987
test	0,0067		30,506		44,08		35,12		1,03		0,0156		1,09		0,02		0,005	
mnist_784																		
train	39,916	0,967	838,573	0,967	717,7	0,967	450,52	0,967	2,49	0,948	5,5304	0,9672	5,37	0,948	0,77	0,903	5,65	0,945
test	0,03		64,767		50,13		68,45		1,27		0,0156		1,32		0,03		0,005	
covtype_2vr																		
train	>40000	-	>180000	-	>3600	-	>40000	-	36,39	0,755	238,44	0.545	0,79	0,612	185	0,505	987	0,6347
test	-		-		-		-		0,69		0,1718		0,8		0,17		0,013	
ijcnn1 (with standardization)																		
train	197,21	0,957	200,8	0,957	122,5	0,917	25,96	0,930	0,14	0,602	5,486	0,916	0,08	0,602	7,44	0,905	2,11	0,9232
test	102,73		82,1		1,13		108,78		0,59		0,018		0,58		0,03		0,003	
mnist_576 (with standardization)																		
train	33072,2	0,903	31946,4	0,992	>3600	-	144,98	0,999	73,09	0,822	32,067	0,9905	36,95	0,822	118	0,988	5,31	0,9879
test	1189,63		16,27		-		52,115		1,57		0,047		1,57		0,03		0,004	
mnist_784 (with standardization)																		
train	31083,3	0,903	19446,2	0,903	>3600	-	483,86	0,995	13,98	0,678	62,699	0,9458	13,26	0,679	83,8	0,947	6,256	0,9637
test	1039,41		730,48		-		82,088		1,32		0,0312		1,3		0,02		0,006	
covtype_2vr (with standardization)																		
train	>40000	-	>180000	-	>3600	-	>40000	-	2,72	0,756	157,99	0,7561	2,47	0,754	130	0,693	11,61	0,7534
test	-		-		-		-		4,24		0,0937		4,29		0,08		0,02	

Data description

Set	Objects on train	Objects on recognition	Number of features
ijcnn1	35000	91701	22
mnist-576-rbf-8vr	60000	10000	576
mnist-784-poly-8vr	60000	10000	784
covtype-2vr	300000	281012	54

Characteristics of CS

*The supercomputer complex of MSU “Lomonosov”: 1 node: Intel Xeon X5570 (2.93 GHz), 8 core, 8Gb RAM, node: 5104

** PC: Intel® Core™ i5-4210U (2.4GHz), 2 core, 6Gb RAM

Results of testing on real data

	SVMlight* (C)		libSVM* (C)		SVC** (LibSVM, python)		πSVM* (1 process)		liblinear* (C)		linear-SVC** (Liblinear, python)		MPI - liblinear* (1 process)		SGD** (python)		proposed method**	
	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy
ijcnn1																		
train	12,41	0,917	22,193	0,917	40,4	0,917	13,99	0,917	0,16	0,913	0,9151	0,9164	0,45	0,913	9,89	0,905	0,412	0,9147
test	0,0267		21,146		25,72		24,67		0,34		0,02		0,09		0,01		0,004	
mnist_576																		
train	51,26	0,994	512,736	0,994	499,8	0,994	95,41	0,994	2,17	0,989	4,5533	0,9935	4,73	0,992	0,51	0,903	5,74	0,987
test	0,0067		30,506		44,08		35,12		1,03		0,0156		1,09		0,02		0,005	
mnist_784																		
train	39,916	0,967	838,573	0,967	717,7	0,967	450,52	0,967	2,49	0,948	5,5304	0,9672	5,37	0,948	0,77	0,903	5,65	0,945
test	0,03		64,767		50,13		68,45		1,27		0,0156		1,32		0,03		0,005	
covtype_2vr																		
train	>40000	-	>180000	-	>3600	-	>40000	-	36,39	0,755	238,44	0.545	0,79	0,612	185	0,505	987	0,6347
test	-		-		-		-		0,69		0,1718		0,8		0,17		0,013	
ijcnn1 (with standardization)																		
train	197,21	0,957	200,8	0,957	122,5	0,917	25,96	0,930	0,14	0,602	5,486	0,916	0,08	0,602	7,44	0,905	2,11	0,9232
test	102,73		82,1		1,13		108,78		0,59		0,018		0,58		0,03		0,003	
mnist_576 (with standardization)																		
train	33072,2	0,903	31946,4	0,992	>3600	-	144,98	0,999	73,09	0,822	32,067	0,9905	36,95	0,822	118	0,988	5,31	0,9879
test	1189,63		16,27		-		52,115		1,57		0,047		1,57		0,03		0,004	
mnist_784 (with standardization)																		
train	31083,3	0,903	19446,2	0,903	>3600	-	483,86	0,995	13,98	0,678	62,699	0,9458	13,26	0,679	83,8	0,947	6,256	0,9637
test	1039,41		730,48		-		82,088		1,32		0,0312		1,3		0,02		0,006	
covtype_2vr (with standardization)																		
train	>40000	-	>180000	-	>3600	-	>40000	-	2,72	0,756	157,99	0,7561	2,47	0,754	130	0,693	11,61	0,7534
test	-		-		-		-		4,24		0,0937		4,29		0,08		0,02	

Data description

Set	Objects on train	Objects on recognition	Number of features
ijcnn1	35000	91701	22
mnist-576-rbf-8vr	60000	10000	576
mnist-784-poly-8vr	60000	10000	784
covtype-2vr	300000	281012	54

Characteristics of CS

*The supercomputer complex of MSU "Lomonosov":
1 node: Intel Xeon X5570 (2.93 GHz), 8 core,
8Gb RAM, node: 5104
** PC: Intel® Core™ i5-4210U (2.4GHz), 2 core,
6Gb RAM

Results of testing on real data

	SVMlight* (C)		libSVM* (C)		SVC** (LibSVM, python)		πSVM* (1 process)		liblinear* (C)		linear-SVC** (Liblinear, python)		MPI - liblinear* (1 process)		SGD** (python)		proposed method**	
	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy	time (s)	accu-racy
ijcnn1																		
train	12,41	0,917	22,193	0,917	40,4	0,917	13,99	0,917	0,16	0,913	0,9151	0,9164	0,45	0,913	9,89	0,905	0,412	0,9147
test	0,0267		21,146		25,72		24,67		0,34		0,02		0,09		0,01		0,004	
mnist_576																		
train	51,26	0,994	512,736	0,994	499,8	0,994	95,41	0,994	2,17	0,989	4,5533	0,9935	4,73	0,992	0,51	0,903	5,74	0,987
test	0,0067		30,506		44,08		35,12		1,03		0,0156		1,09		0,02		0,005	
mnist_784																		
train	39,916	0,967	838,573	0,967	717,7	0,967	450,52	0,967	2,49	0,948	5,5304	0,9672	5,37	0,948	0,77	0,903	5,65	0,945
test	0,03		64,767		50,13		68,45		1,27		0,0156		1,32		0,03		0,005	
covtype_2vr																		
train	>40000	-	>180000	-	>3600	-	>40000	-	36,39	0,755	238,44	0.545	0,79	0,612	185	0,505	987	0,6347
test	-		-		-		-		0,69		0,1718		0,8		0,17		0,013	
ijcnn1 (with standardization)																		
train	197,21	0,957	200,8	0,957	122,5	0,917	25,96	0,930	0,14	0,602	5,486	0,916	0,08	0,602	7,44	0,905	2,11	0,9232
test	102,73		82,1		1,13		108,78		0,59		0,018		0,58		0,03		0,003	
mnist_576 (with standardization)																		
train	33072,2	0,903	31946,4	0,992	>3600	-	144,98	0,999	73,09	0,822	32,067	0,9905	36,95	0,822	118	0,988	5,31	0,9879
test	1189,63		16,27		-		52,115		1,57		0,047		1,57		0,03		0,004	
mnist_784 (with standardization)																		
train	31083,3	0,903	19446,2	0,903	>3600	-	483,86	0,995	13,98	0,678	62,699	0,9458	13,26	0,679	83,8	0,947	6,256	0,9637
test	1039,41		730,48		-		82,088		1,32		0,0312		1,3		0,02		0,006	
covtype_2vr (with standardization)																		
train	>40000	-	>180000	-	>3600	-	>40000	-	2,72	0,756	157,99	0,7561	2,47	0,754	130	0,693	11,61	0,7534
test	-		-		-		-		4,24		0,0937		4,29		0,08		0,02	

Data description

Set	Objects on train	Objects on recognition	Number of features
ijcnn1	35000	91701	22
mnist-576-rbf-8vr	60000	10000	576
mnist-784-poly-8vr	60000	10000	784
covtype-2vr	300000	281012	54

Characteristics of CS

*The supercomputer complex of MSU “Lomonosov”: 1 node: Intel Xeon X5570 (2.93 GHz), 8 core, 8Gb RAM, node: 5104

** PC: Intel® Core™ i5-4210U (2.4GHz), 2 core, 6Gb RAM

Thank you for your attention!