Data sets	ER (ours)	EP	5-CV	LOO	CKTA	FSM
			$\lambda = 0.01$			
australian	14.01±2.10	14.20±1.95	14.78±1.87	14.30±1.85	**13.98±2.10**	44.15±3.30
heart	17.08±3.28	18.56±3.59	16.91±3.64	**16.71±3.29**	16.75±3.50	44.65±4.35
ionosphere	8.98±2.39	9.65±2.58	**5.97±1.58**	6.32±2.33	32.89±8.71	35.68±3.13
breast-cancer	**3.59±0.97**	3.72±1.04	3.77±1.09	3.61±0.94	31.72±9.40	4.36±0.93
diabetes	22.30±2.30	22.81±2.67	22.43±2.69	**22.06±2.31**	35.09±2.87	35.09±2.87
german.numer	**23.76±2.34**	26.14±1.70	24.11±2.16	24.14±2.31	30.36±2.57	30.36±2.57
liver-disorders	**28.56±3.75**	30.80±3.40	29.17±4.24	29.04±4.37	36.83±6.53	40.83±4.72
a2a	17.74±0.95	19.22±1.07	17.80±1.03	**17.67±0.90**	25.21±1.40	25.21±1.40
			$\lambda = 0.1$			
australian	**13.46±1.91**	13.59±2.03	14.59±2.07	14.41±1.73	19.89±6.18	44.15±3.30
heart	**16.87±3.39**	17.12±3.40	17.08±3.82	17.16±3.81	40.29±10.03	44.65±4.35
ionosphere	7.87±2.36	8.89±2.57	**5.30±1.29**	6.38±1.92	35.75±3.04	35.68±3.13
breast-cancer	3.53±0.97	4.23±1.09	3.79±0.91	**3.51±0.94**	32.70±6.57	3.84±0.81
diabetes	22.25±2.53	22.99±2.75	22.35±2.40	**22.16±2.35**	35.09±2.87	35.09±2.87
german.numer	**23.77±2.37**	23.89±2.13	24.08±2.15	23.99±2.31	30.36±2.57	30.36±2.57
liver-disorders	**29.01±4.19**	34.20±3.93	30.19±4.13	30.77±4.00	35.54±7.29	40.83±4.72
a2a	17.81±1.19	18.11±1.22	17.92±1.01	**17.68±0.97**	25.21±1.40	25.21±1.40
			$\lambda = 1$			
australian	**14.12±1.62**	14.17±1.58	14.67±2.07	14.61±1.86	44.07±3.35	44.15±3.30
heart	**16.83±3.56**	17.12±3.30	17.08±3.01	17.12±3.72	44.65±4.35	44.65±4.35
ionosphere	7.81±2.60	11.02±3.13	**5.43±1.97**	6.57±2.08	35.75±3.04	35.68±3.13
breast-cancer	3.43±0.95	4.54±1.55	3.53±0.83	**3.25±0.94**	34.44±2.32	3.54±0.87
diabetes	22.39±2.53	24.65±3.29	22.48±2.22	**22.19±2.20**	35.09±2.87	35.09±2.87
german.numer	24.57±2.38	24.59±2.47	24.24±2.41	**24.13±2.37**	30.36±2.57	30.36±2.57
liver-disorders	**29.46±3.69**	41.96±4.45	30.38±3.99	30.87±3.42	36.19±6.09	40.83±4.72
a2a	18.17±1.23	18.66±1.31	18.01±1.00	**17.83±1.08**	25.21±1.40	25.21±1.40
			$\lambda = 10$			
australian	**13.85±1.88**	14.19±1.71	14.52±1.73	14.35±2.06	44.07±3.35	44.07±3.35
heart	17.04±3.62	17.37±3.59	**16.79±3.44**	17.04±3.36	44.65±4.35	44.65±4.35
ionosphere	8.89±2.63	14.22±3.74	6.92±1.87	**6.76±1.88**	36.06±3.06	36.06±3.06
breast-cancer	3.24±0.89	5.71±1.59	3.37±0.93	**3.22±1.08**	35.04±2.45	3.61±0.88
diabetes	23.41±3.07	34.99±3.13	23.09±2.78	**22.54±2.51**	35.09±2.87	35.09±2.87
german.numer	27.13±2.65	29.64±3.30	26.94±2.80	**26.86±2.87**	30.36±2.57	30.36±2.57
liver-disorders	**34.97±4.47**	41.89±4.39	35.35±4.46	35.13±4.49	39.26±4.74	41.86±4.46
a2a	18.79±1.19	19.26±1.60	18.77±1.21	**18.71±1.21**	25.23±1.43	25.23±1.43

Table 1: Comparison of mean test errors between our eigenvalues ratio criterion (ER) and other ones including 5-CV, LOO, CKTA, FSM and EP. We bold the numbers of the best method, and underline the numbers of the other methods which are not significantly worse than the best one.

where $C_4 = 12hBk\sqrt{2\kappa}$, $C_5 = 96h^2Bkt + (96k + 22M + 5Bk)\log(3/\delta)$, $h = 2(M + D)$, $C_6 = (D + M)^2/(k - 1)$, $B = (D + M)^2$.

The idea of proof is same as that of Theorem 2.

The convergence rate of KRR is $O\left(1/\sqrt{n\beta_t} + 1/n\right)$. Under the assumption of algebraically decreasing eigenvalues, the converges rate can reach $O(1/n)$. This theorem also indicates that the kernel with small $R_{\mathrm{emp}}(S)$ and $1/\beta_t$ can guarantee good generalization performance.

Least Squares Support Vector Machine (LSSVM)

LSSVM is a popular classifier which has the same loss function as that of KRR. Thus, applying Theorem 3 with $M = 1$, we can obtain the following corollary:

Corollary 1. *If the t-eigenvalues ratio of K is β_t, then for LSSVM, with probability at least $1 - \delta$, for any $k > 1$,*

$$R(S) \leq R_{\mathrm{emp}}(S) + C_7\sqrt{1/(n\beta_t)} + C_8/n + C_9,$$

where $C_7 = 12hk\sqrt{2\kappa}$, $C_8 = 96Bh^2kt + (96k + 22 + 5Bk)\log(3/\delta)$, $h = 2(1 + D)$, $C_9 = (1 + D)^2/(k - 1)$, $B = (D + 1)^2$.

Kernel Selection with Eigenvalues Ratio

In this section, we will present a novel kernel selection criterion with ER to guarantee good generalization performance.

From the generalization error bounds derived in above section, to guarantee good generalization performance, we can choose the kernel function by minimizing $R_{\mathrm{emp}}(S)$ and $1/\beta_t$. Thus, we apply the following eigenvalues ratio criterion for kernel selection:

$$\operatorname*{arg\,min}_{K \in \mathcal{K}} R_{\mathrm{emp}}(S) + \eta \cdot n/\beta_t =: KS(K)$$