
The 16th ACM SIGKDD International Conference
on Knowledge Discovery and Data Mining

# Learning with Cost Intervals

Xu-Ying Liu and Zhi-Hua Zhou

LAMDA Group

National Key Laboratory for Novel Software Technology

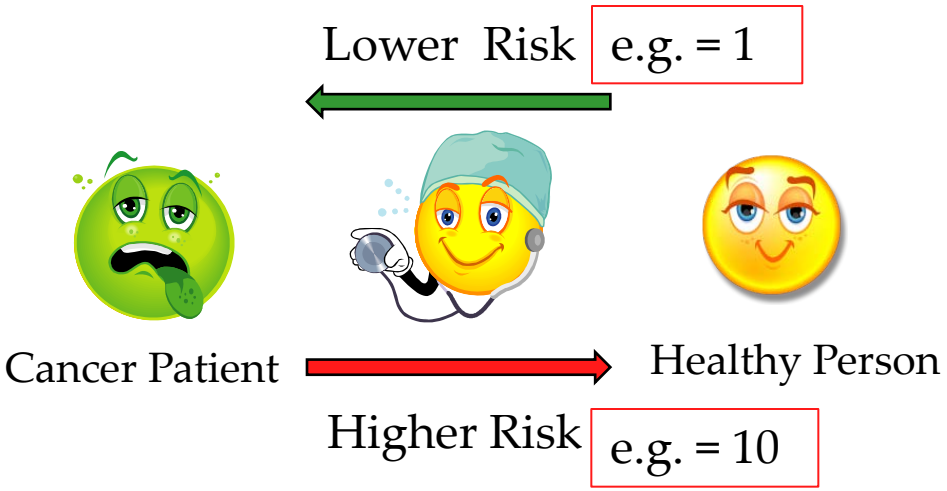
Nanjing University, China

{liuxy, zhouzh}@lamda.nju.edu.cn

# Imprecise Misclassification Cost



In many real applications, different mistakes often have different costs. Learning methods should minimize the total cost instead of simply minimizing the error rate




Existing cost-sensitive learning methods assume that precise cost values are given

**It is often difficult to get precise cost values**

... What can we do ?

# Outline



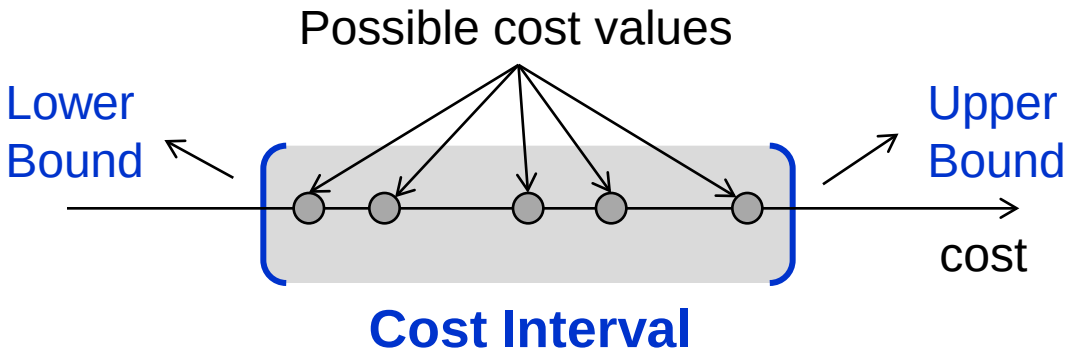
- Introduction
- Our Method
- Experiments
- Conclusion

# Cost Interval



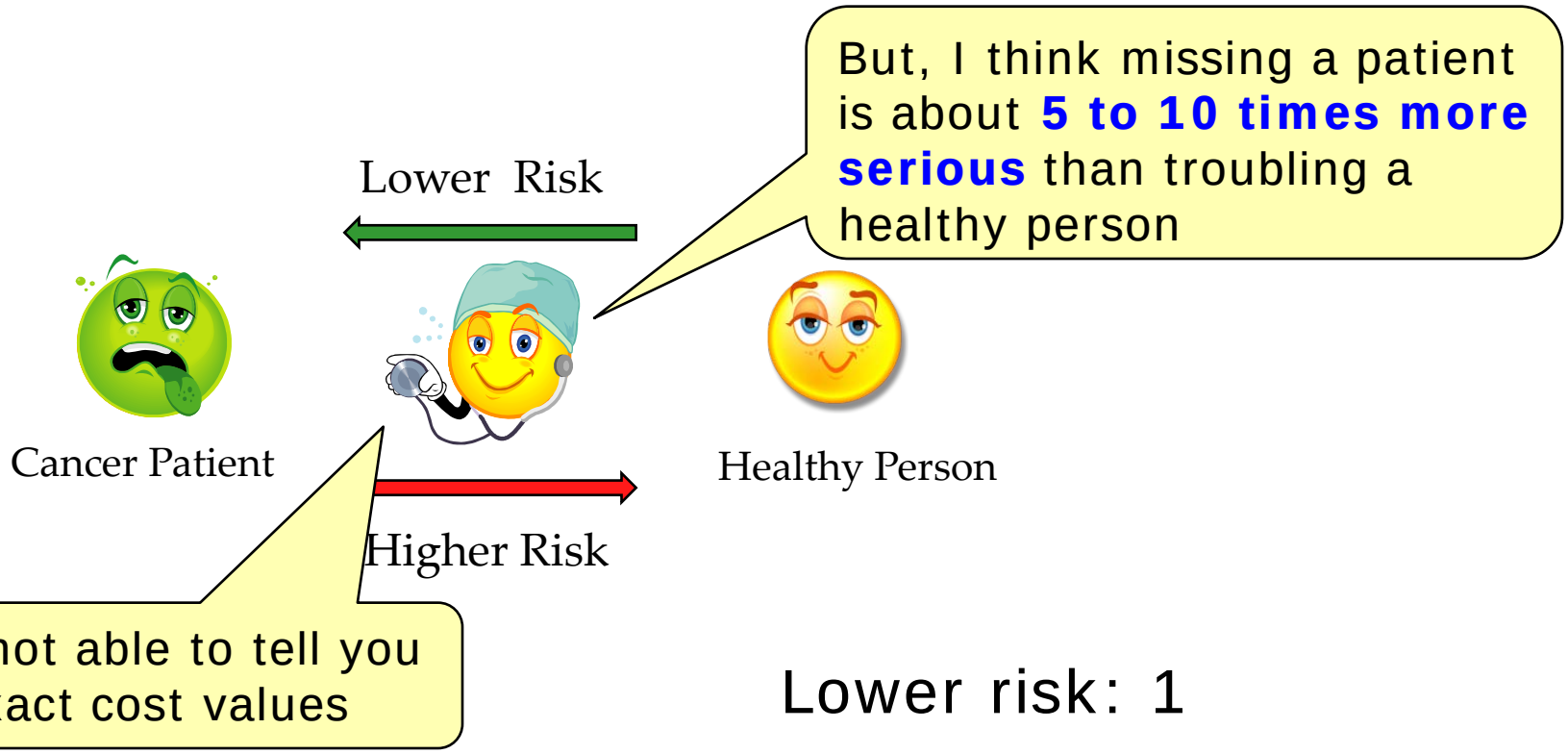
Although the user knows that one type of mistake is more severe than another type, it is often difficult to provide precise cost values

However, obtaining a **cost interval** is feasible in most cases



# Cost Interval (cont.)






Lower risk: 1

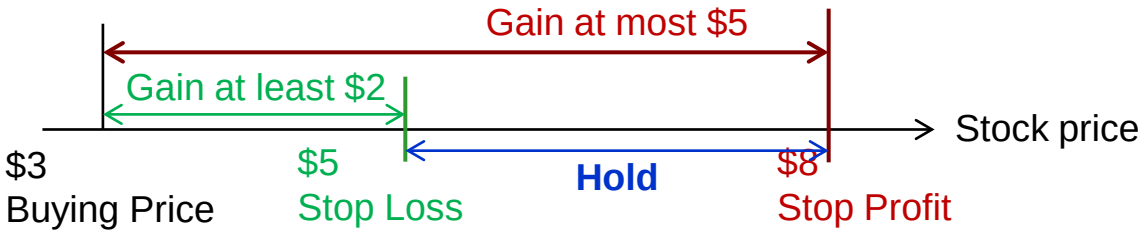
Higher risk: [5, 10]

# Cost Interval (cont.)



Possible ways to obtain cost intervals:

- Natural upper and lower bounds




- Transforming from confidence intervals

Breast Cancer Report: the risk for breast cancer is greater for women with ADH of (IRR=5.0; 95% CI=2.26-11.0)

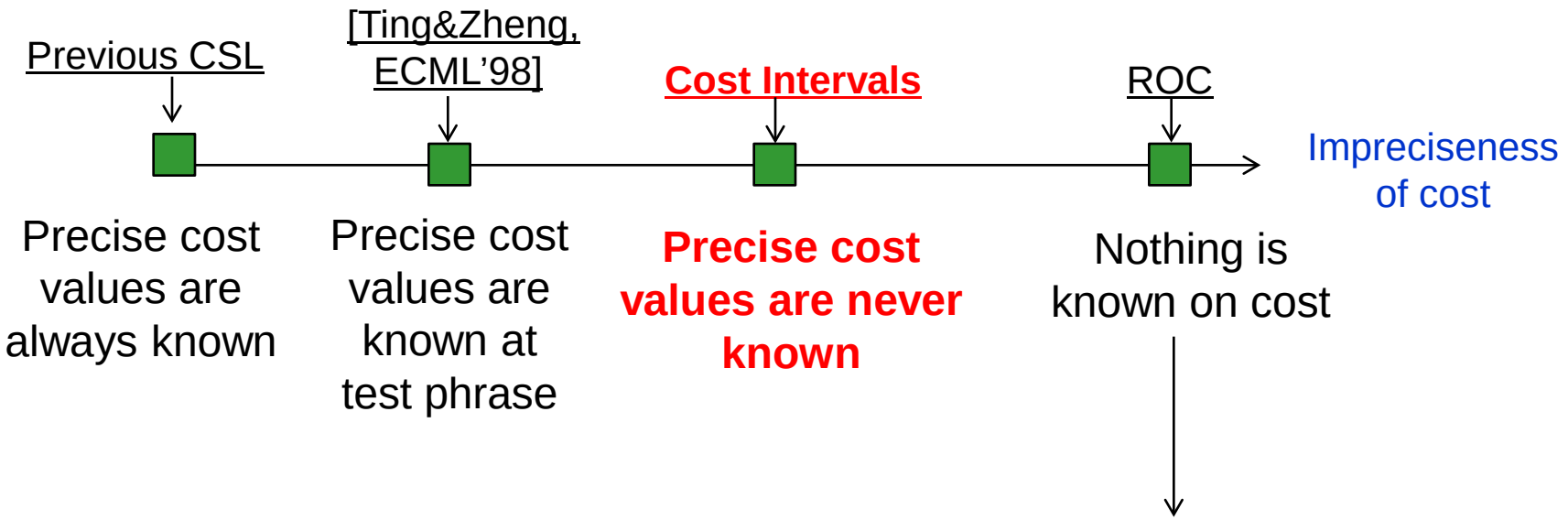
Incidence rate ratio is about ↗ ↖ confidence interval

- Expert opinions
- . . .

# Related Work



From the view of Cost Sensitive Learning:




if used directly to deal with costs, it assume that: Costs vary in $[-\infty, +\infty]$ , and nothing is known about costs (including which class has higher cost ! )

# Outline



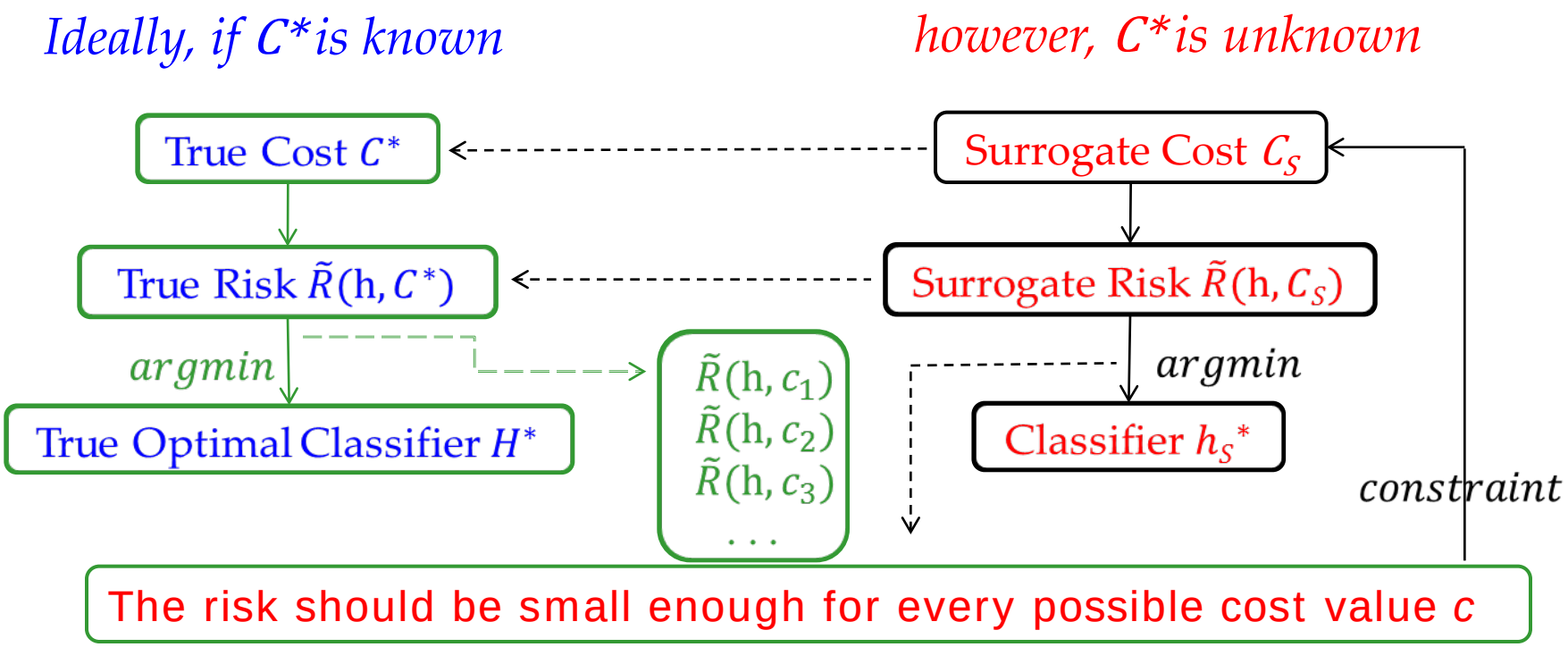
- Introduction
- Our Method
- Experiments
- Conclusion

# Basic Idea



$c_- = 1,\ c_+ = c \in [C_{min}, C_{max}]$

true cost $C^*$: the true value of $c$



# Basic Idea



$c_- = 1,\ c_+ = c \in [C_{min}, C_{max}]$

true cost $C^*$: the true value of $c$

$$\min_{h, C_s} \quad \tilde{R}(h, C_s)$$

$$\text{s.t.} \quad p(\tilde{R}(h, c) < \epsilon) > 1 - \delta,\ \forall c \in [C_{min}, C_{max}]$$

$$C_{min} \le C_s \le C_{max}.$$

*however, C*is unknown*

Surrogate Cost $C_s$

Surrogate Risk $\tilde{R}(\mathrm{h}, C_s)$

*argmin*

Classifier $h_s{}^*$

$\tilde{R}(\mathrm{h}, c_1)$
$\tilde{R}(\mathrm{h}, c_2)$
$\tilde{R}(\mathrm{h}, c_3)$
...

*constraint*

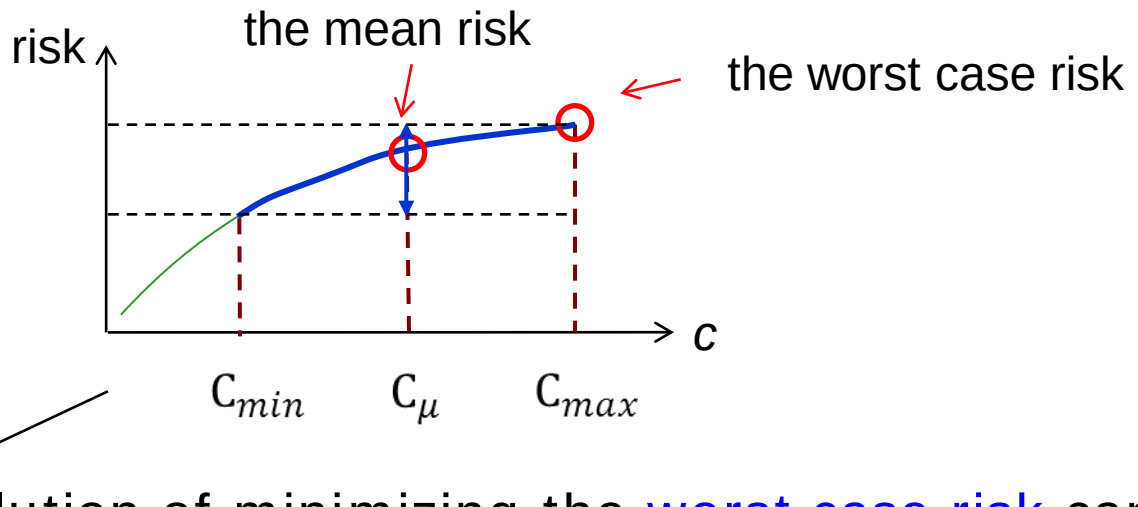
**However, infinite number of constraints !**

The risk should be small enough for every possible cost value $c$

# Relaxation



Try to solve a relaxation with a small number of informative constraints




- The optimal solution of minimizing the worst case risk can make all the constraints in the original formulation hold
- Minimizing the mean risk will reduce the overall distortion from the true risk

If cost distribution is known: $c \sim v$

$\min \int_c risk(c)\, dv$

**Minimize the worst case risk as well as the mean risk**

# The CISVM Method



## CISVM: Cost-Interval Sensitive SVM

Loss Function: $L(C_p, f(x), y) = I_{y=+}[C_p - yf(x)]_+ + I_{y=-}[1 - yf(x)]_+$

Loss of "+" Loss of "-"

- Step 1. Minimize the worst case risk: train SVM

$$\begin{aligned} \min_{w,b,\xi \geq 0} \quad & \|w\|^2/2 + \lambda \sum_{i=1}^{n} \xi_i \qquad (8) \\ \text{s.t.} \quad & y_i(w^{\mathrm{T}}\phi(x_i) + b) \geq C_{max} - \xi_i, \ \forall i : y_i = +1 \\ & y_i(w^{\mathrm{T}}\phi(x_i) + b) \geq 1 - \xi_i, \ \forall i : y_i = -1 \end{aligned}$$

- Step 2. Minimize the mean risk: parameter selection of ($\lambda$,$K$)

To guarantee that all constrains in the original formulation will hold, minimizing the worst case risk is put as the first target

# CSSVM



**CSSVM: Cost-Sensitive SVM** [Brefeld et al, ECML'03]
which was designed for precise cost values

Loss Function: $L_{CS}(C_p, f(x), y) = I_{y=+}\boxed{C_p[1 - yf(x)]_+} + I_{y=-}\boxed{[1 - yf]_+}$,

Loss of "+" Loss of "-"

Where $C_p$ is the precise cost value given in advance

Intuitively, given a cost interval, maybe we can apply CSSVM by assuming the true cost as:

- ✓ The upper bound of the interval -> **CSMax**
- ✓ The lower bound of the interval -> **CSMin**
- ✓ The mean of the interval -> **CSMean**

# Some Theoretical Results



**CISVM :** THEOREM 1. *Suppose $C_p$ is a random value in interval $[C_{min}$, $C_{max}]$, the loss function $L(C_p)$ is achieved when $C_p = C_{max}$:*

$$\tilde{R}_L(C_{max}) = \sup_{C_p} \tilde{R}_L(C_p). \quad (6)$$

*since $d \cdot n_+$ is a constant.*

$L(C_{max})$ is the worst case risk

The optimal solutions of minimizing the worst case risk and the mean risk are different

**CSMax :** *... is achieved when $C_p = C_{max}$:*

CSMax minimizes the worst case risk

**CSMin :** *... $L_{CS}(C_p)$ ... $L_{CS}(C_{min})$.*

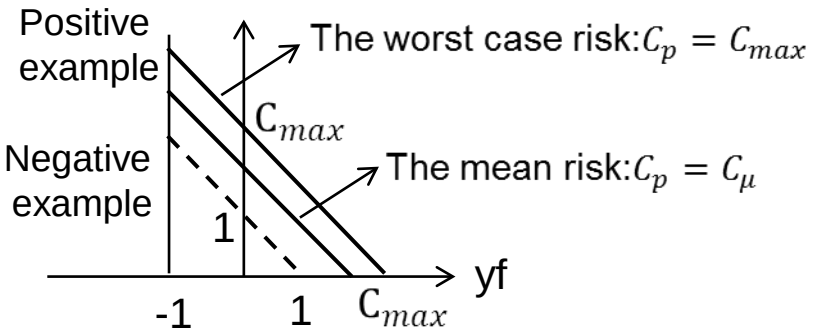
CSMin minimizes the best case risk

**CSMean :**

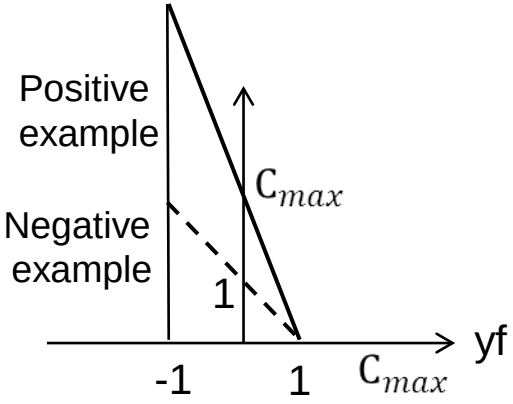
CSMean minimizes the mean risk. The optimal solution of CSMean is different from that of CSMax and CSMin

# Loss Functions

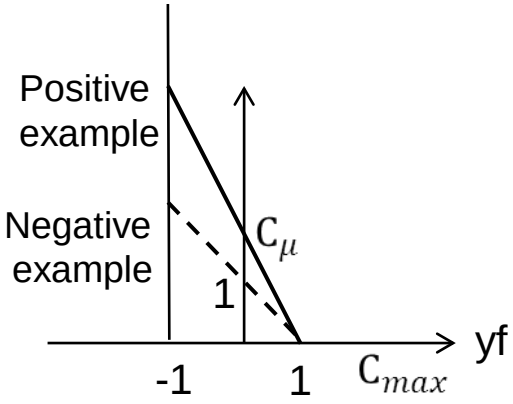




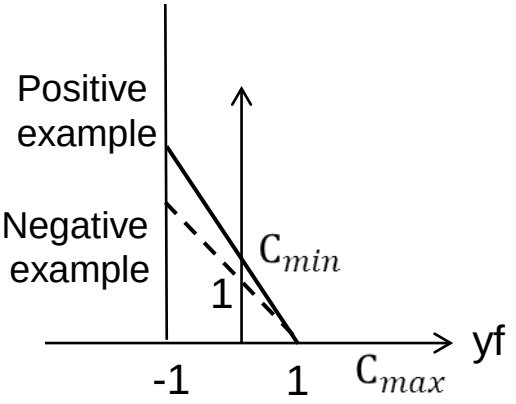

**CISVM:** $L = I(y == -1)[1 - yf]_+ + I(y == 1)[C_p - yf]_+$




**CSMax:** $L_{cs}(C_{\max})$




**CSMean:** $L_{cs}(C_\mu)$




**CSMin:** $L_{cs}(C_{\min})$

$$L_{cs}(C_p) = I(y == -1)[1 - yf]_+ + I(y == 1)C_p[1 - yf]_+$$

# Robust



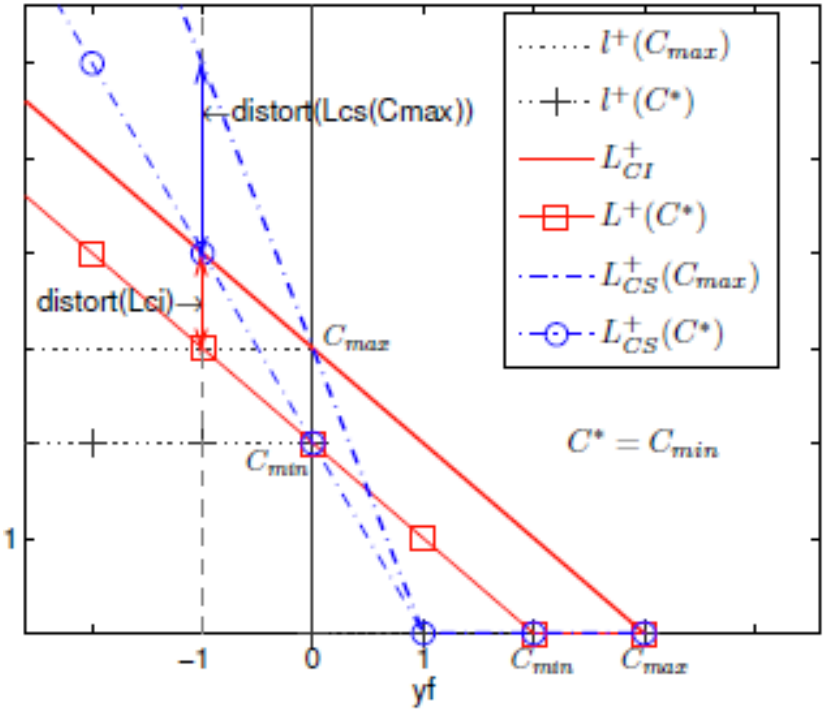
$$\mathrm{distort}(\mathcal{L}(C_p)) = \sup_{C^*,h(x),y} |\mathcal{L}(C^*, h(x), y) - \mathcal{L}(C_p, h(x), y)|.$$

$$\mathrm{robust}(\mathcal{L}(C_p)) = \frac{\inf_{C_p} \mathrm{distort}(l(C_p))}{\mathrm{distort}(\mathcal{L}(C_p))}$$

**The smaller the maximal distortion from the true risk, the better**

CSMax: distort = 4d, robust = 0.25
CSMean: distort = 2d, robust = 0.5
CSMin: distort = 4d, robust = 0.25
CISVM: distort = d, robust = 1

$(d = 0.5(C_{max} - C_{min}))$



# Utilizing Cost Distribution



If we know more information, we can do better

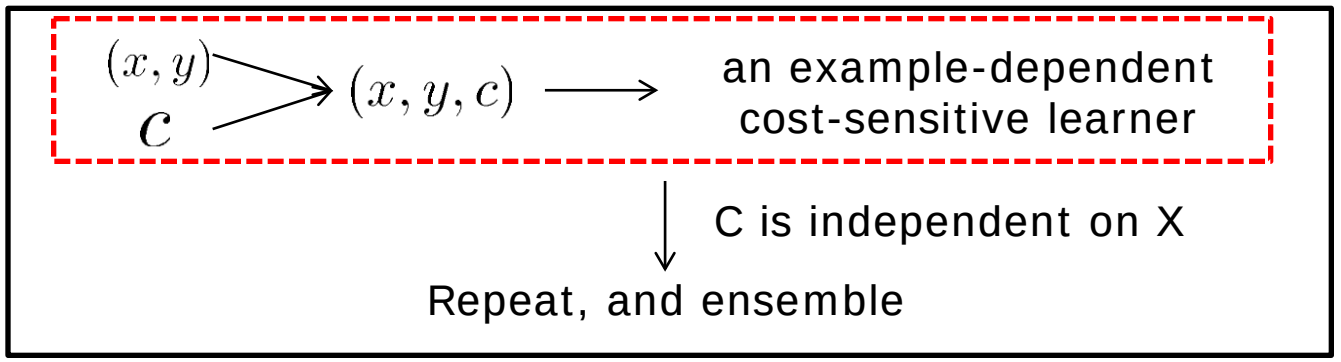
e.g., if we know the cost distribution:

- Cost $c$ drawn *i.i.d.* from distribution $v(c)$
- Goal: To minimize $R_{CD}(h, v) = E_{c \sim v}[R(h, c)]$

The **CODIS** (COst DIStribution) method:

By Theorem 3, we get $E_{c \sim \mathcal{C}}[R(h(g(c)), c)] = E_{(x,y,c) \sim \mathcal{D}}[l(c, h(x, g(c)), y)]$.

Thus,



# Outline



- Introduction
- Our Method
- Experiments
- Conclusion

# Settings



**Compared methods:**

CISVM, CSMin, CSMean, CSMax, SVM (cost-blind)

All with RBF kernel

**Data:**

15 UCI data sets; switching +/-; 25 cost intervals

Overall **750** tasks

Table 1: Cost Intervals

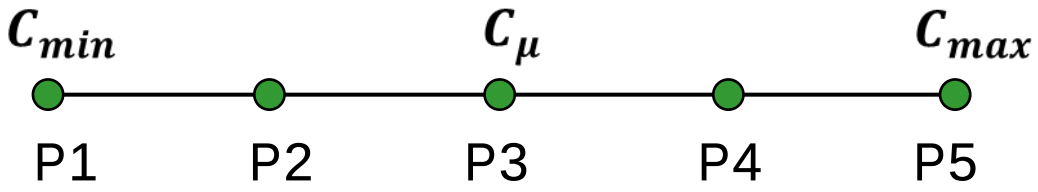
| Width | Counts | Cost Intervals |
|---|---|---|
| 3 | 10 | $[1,3],[4,6],[7,9],[10,12],[13,15]$ $[16,18],[19,21],[22,24],[25,27],[28,30]$ |
| 5 | 6 | $[1,5],[6,10],[11,15],[16,20],[21,25],[26,30]$ |
| 11 | 5 | $[1,11],[5,15],[10,20],[15,25],[20,30]$ |
| 15 | 4 | $[1,16],[5,20],[10,25],[15,30]$ |

# Evaluation



30 times hold-out tests, 2/3 training, 1/3 testing

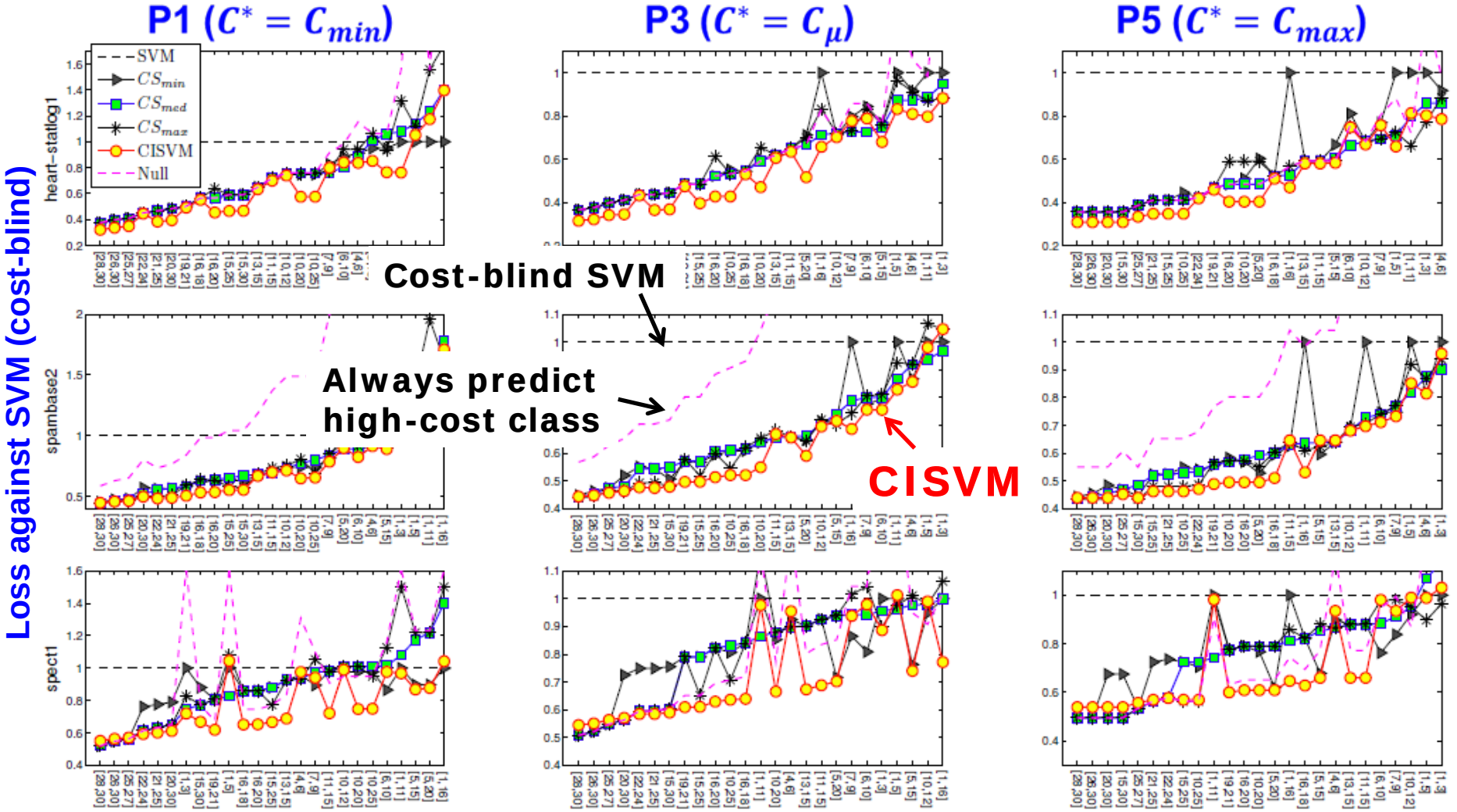
Quadrisection points of the cost interval are considered in turn as the true cost



# Results

**the lower, the better**

LAMDA
Learning And Mining from DatA
http://lamda.nju.edu.cn



# Results (cont.)

| | | SVM | CSMIN | CSMEAN | CSMAX |
|---|---|---|---|---|---|
| P1 | CSMIN | 457/271/22 | | | |
| | CSMEAN | 484/162/104 | *83/487/180* | | |
| | CSMAX | 488/160/102 | *107/461/182* | *92/522/136* | |
| | CISVM | 521/96/133 | 257/262/231 | 316/248/186 | 303/259/188 |
| P2 | CSMIN | 464/272/14 | | | |
| | CSMEAN | 533/175/42 | 121/514/115 | | |
| | CSMAX | 525/186/39 | 139/494/117 | *82/542/126* | |
| | CISVM | 578/97/75 | 304/263/183 | 296/269/185 | 282/282/186 |
| P3 | CSMIN | 466/273/11 | | | |
| | CSMEAN | 550/180/20 | 143/509/98 | | |
| | CSMAX | 554/173/23 | 159/495/96 | *77/560/113* | |
| | CISVM | 593/93/64 | 304/271/175 | 276/292/182 | 260/309/181 |
| P4 | CSMIN | 472/270/8 | | | |
| | CSMEAN | 554/180/16 | 150/503/97 | | |
| | CSMAX | 561/176/13 | 168/493/89 | 89/565/96 | |
| | CISVM | 599/91/60 | 299/275/176 | 256/296/198 | 232/327/191 |
| P5 | CSMIN | 475/268/7 | | | |
| | CSMEAN | 556/178/16 | 150/503/97 | | |
| | CSMAX | 566/170/14 | 173/488/89 | 100/556/94 | |
| | CISVM | 606/85/59 | 299/270/181 | 252/294/204 | 226/325/199 |
| All | CSMIN | 2334/1354/62 | | | |
| | CSMEAN | 2677/875/198 | 647/2516/587 | | |
| | CSMAX | 2694/865/191 | 746/2431/573 | *440/2745/565* | |
| | CISVM | 2897/462/391 | 1463/1341/946 | 1396/1399/955 | 1303/1502/945 |

**Row vs. column**

**w/t/l counts: after t-tests (95% SI)**

**Bold: sign-tests (95% SI)**

All cost-sensitive methods are significantly better than cost-blind

**CISVM is significantly better than the other methods, except that it is comparable with CSMin when the true cost is P1, and CSMax when the true cost is P4 and P5**

# Results (cont.)

M1: cost-blind SVM
M2: CSMin
M3: CSMean
M4: CSMax
M5: CISVM

LAMDA
Learning And Mining from DatA
http://lamda.nju.edu.cn

**Counts of SBPs (Significantly Best Performances),** t-tests 95% SI

| | P1 | | | | | P2 | | | | | P3 | | | | | P4 | | | | | P5 | | | | | All | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | M1 | M2 | M3 | M4 | M5 | M1 | M2 | M3 | M4 | M5 | M1 | M2 | M3 | M4 | M5 | M1 | M2 | M3 | M4 | M5 | M1 | M2 | M3 | M4 | M5 | M1 | M2 | M3 | M4 | M5 |
| [1, 3] | **20** | **20** | 10 | 8 | 11 | 10 | 10 | 14 | 12 | **16** | 7 | 7 | 15 | 14 | **20** | 7 | 7 | 17 | 17 | **19** | 7 | 7 | 16 | **18** | 17 | 51 | 51 | 72 | 69 | **83** |
| [4, 6] | 7 | 19 | 14 | 14 | **21** | 7 | 19 | 15 | 14 | **21** | 7 | 19 | 15 | 16 | **21** | 7 | **20** | 16 | 16 | **20** | 6 | **21** | 16 | 17 | **21** | 34 | 98 | 76 | 77 | **104** |
| [7, 9] | 7 | 18 | 14 | 13 | **22** | 7 | 18 | 14 | 13 | **21** | 7 | 18 | 15 | 13 | **21** | 7 | 19 | 17 | 14 | **20** | 7 | 19 | 18 | 14 | **20** | 35 | 92 | 78 | 67 | **104** |
| [10, 12] | 7 | 18 | 16 | 14 | **21** | 7 | 19 | 16 | 15 | **22** | 7 | 19 | 18 | 16 | **22** | 7 | 18 | 18 | 16 | **21** | 7 | 18 | 19 | 17 | **21** | 35 | 92 | 87 | 78 | **107** |
| [13, 15] | 6 | 18 | 18 | 18 | **22** | 5 | 17 | 18 | 18 | **22** | 5 | 18 | 18 | 20 | **22** | 5 | 18 | 18 | 20 | **22** | 5 | 19 | 17 | 19 | **22** | 26 | 90 | 89 | 95 | **110** |
| [16, 18] | 4 | 17 | 16 | 14 | **22** | 4 | 17 | 16 | 14 | **22** | 4 | 17 | 16 | 14 | **22** | 4 | 18 | 17 | 15 | **22** | 4 | 19 | 17 | 16 | **22** | 20 | 88 | 82 | 73 | **110** |
| [19, 21] | 4 | 15 | 14 | 16 | **21** | 4 | 15 | 14 | 16 | **21** | 3 | 15 | 14 | 16 | **21** | 3 | 15 | 14 | 16 | **21** | 3 | 15 | 14 | 16 | **21** | 17 | 75 | 70 | 80 | **105** |
| [22, 24] | 4 | 16 | 19 | 19 | **23** | 4 | 17 | 20 | 20 | **23** | 4 | 17 | 20 | 20 | **23** | 4 | 18 | 21 | 21 | **23** | 4 | 18 | 21 | 21 | **23** | 20 | 86 | 101 | 101 | **115** |
| [25, 27] | 3 | 21 | 20 | 22 | **23** | 3 | 21 | 20 | **22** | **22** | 3 | 21 | 20 | 21 | **22** | 3 | 21 | 20 | **22** | **22** | 3 | 21 | 20 | 22 | **23** | 15 | 105 | 100 | 109 | **112** |
| [28, 30] | 3 | **24** | **24** | 23 | 21 | 3 | **24** | **24** | 23 | 21 | 3 | **24** | **24** | 23 | 21 | 3 | **24** | **24** | 23 | 21 | 3 | **24** | **24** | 23 | 21 | 15 | **120** | **120** | 115 | 105 |
| [1, 5] | **18** | **18** | 8 | 6 | 8 | 11 | 11 | 14 | 13 | **17** | 7 | 7 | 18 | 15 | **19** | 7 | 7 | **20** | 15 | 18 | 8 | 8 | **20** | 15 | 17 | 51 | 51 | **80** | 64 | 79 |
| [6, 10] | 6 | 15 | 14 | 11 | **21** | 6 | 17 | 15 | 12 | **20** | 6 | 17 | 16 | 12 | **18** | 6 | 18 | 17 | 13 | **19** | 6 | **19** | 17 | 14 | **19** | 30 | 86 | 79 | 62 | **97** |
| [11, 15] | 5 | 15 | 13 | 13 | **21** | 5 | 16 | 13 | 14 | **21** | 5 | 18 | 17 | 18 | **21** | 5 | 19 | 18 | 20 | **21** | 4 | 19 | 19 | 20 | **21** | 24 | 87 | 80 | 85 | **105** |
| [16, 20] | 6 | 17 | 17 | 16 | **22** | 4 | 16 | 15 | 16 | **22** | 3 | 17 | 15 | 16 | **22** | 3 | 16 | 15 | 16 | **21** | 3 | 16 | 15 | 16 | **21** | 19 | 82 | 77 | 80 | **108** |
| [21, 25] | 4 | 16 | 17 | 19 | **23** | 4 | 15 | 17 | 19 | **23** | 4 | 16 | 18 | 19 | **23** | 4 | 17 | 19 | 19 | **23** | 3 | 17 | 18 | 20 | **22** | 19 | 81 | 89 | 96 | **114** |
| [26, 30] | 3 | 22 | 22 | **23** | 21 | 3 | 22 | **23** | **23** | 21 | 3 | 22 | **23** | **23** | 21 | 3 | 21 | **23** | **23** | 21 | 3 | 20 | **23** | **23** | 21 | 15 | 107 | 114 | **115** | 105 |
| [1, 11] | **22** | **22** | 7 | 5 | 10 | 8 | 8 | 18 | 9 | **21** | 6 | 6 | 18 | 14 | **20** | 6 | 6 | **21** | 18 | 16 | 6 | 6 | **21** | 20 | 15 | 48 | 48 | **85** | 66 | 82 |
| [5, 15] | 7 | 15 | 12 | 12 | **20** | 7 | 17 | 12 | 14 | **20** | 7 | 20 | 16 | 15 | **24** | 6 | 19 | 19 | 16 | **23** | 6 | 19 | 19 | 18 | **23** | 33 | 90 | 78 | 75 | **110** |
| [10, 20] | 5 | 12 | 12 | 12 | **22** | 5 | 13 | 13 | 12 | **22** | 5 | 15 | 16 | 16 | **22** | 3 | 15 | 16 | 15 | **21** | 3 | 15 | 16 | 15 | **20** | 21 | 70 | 73 | 70 | **107** |
| [15, 25] | 5 | 16 | 14 | 19 | **22** | 5 | 16 | 15 | 20 | **22** | 3 | 15 | 15 | 20 | **21** | 3 | 17 | 16 | 20 | **21** | 3 | 18 | 17 | 21 | **22** | 19 | 82 | 77 | 100 | **108** |
| [20, 30] | 4 | 14 | 19 | 20 | **21** | 3 | 15 | 20 | **21** | **21** | 3 | 16 | 19 | **22** | 21 | 3 | 17 | 23 | **24** | 21 | 3 | 17 | 23 | **24** | 21 | 16 | 79 | 104 | **111** | 105 |
| [1, 16] | **26** | **26** | 8 | 5 | 6 | 7 | 7 | 11 | 13 | **20** | 7 | 7 | 16 | 16 | **22** | 7 | 7 | 18 | 19 | **21** | 6 | 6 | 18 | 18 | **20** | 53 | 53 | 71 | 71 | **89** |
| [5, 20] | 7 | 17 | 14 | 13 | **18** | 7 | 18 | 12 | 14 | **21** | 6 | 16 | 11 | 15 | **22** | 4 | 15 | 14 | 15 | **20** | 4 | 16 | 14 | 14 | **19** | 28 | 82 | 65 | 71 | **100** |
| [10, 25] | 5 | 15 | 13 | 15 | **21** | 5 | 16 | 15 | 18 | **21** | 5 | 18 | 16 | 18 | **20** | 3 | **20** | 15 | 19 | 19 | 3 | **21** | 17 | 19 | **20** | 21 | 90 | 76 | 89 | **101** |
| [15, 30] | 5 | 14 | 15 | 18 | **20** | 4 | 14 | 16 | 19 | **20** | 4 | 17 | 17 | **20** | **20** | 3 | 17 | 18 | **24** | 20 | 3 | 16 | 17 | **22** | 20 | 19 | 78 | 83 | **103** | 100 |
| sum. | 193 | 440 | 370 | 368 | **483** | 138 | 398 | 400 | 404 | **523** | 124 | 402 | 426 | 432 | **531** | 116 | 409 | 454 | 456 | **516** | 113 | 414 | 456 | 462 | **512** | 684 | 2063 | 2106 | 2122 | **2565** |

**CISVM is the best, no matter which one among P1 to P5 is taken as the true cost**

**CISVM is overall the best**

**Using the "true cost" does not always lead to the best performance**

# Results (cont.)



By utilizing information on cost distributions, the performance of **CODIS** is much better

| | SVM | CSSVM | CODIS | r(%) | SVM | CSSVM | CODIS | r(%) | SVM | CSSVM | CODIS | r(%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | U(1,3) | | | | U(1,6) | | | | U(1,11) | | |
| breast-c1 | 31.5±5.0 | 27.6±1.1 | **11.8±2.5** | 57.2 | 46.9±9.0 | 27.7±2.5 | **12.2±2.7** | 55.9 | 106.0±57.2 | 28.0±0.0 | **11.9±2.5** | 57.6 |
| breast-w1 | 10.0±2.6 | 4.9±1.9 | **2.7±1.2** | 45.6 | 10.0±2.6 | 4.9±2.1 | **2.7±1.2** | 45.3 | 10.0±2.6 | 5.4±3.1 | **3.6±1.5** | 33.5 |
| credit-a1 | 41.0±7.0 | 38.6±5.9 | **7.5±2.5** | 80.5 | 63.6±11.7 | 49.7±6.3 | **17.6±4.1** | 64.5 | 113.8±78.1 | 62.9±8.6 | **18.5±4.3** | 70.5 |
| credit-g1 | 81.0±9.7 | 77.9±8.4 | **33.8±5.3** | 56.6 | 253.5±118.8 | 88.3±6.3 | **34.1±5.1** | 61.4 | 410.0±236.0 | 99.1±3.0 | **36.6±5.5** | 63.1 |
| diabetes1 | 64.5±7.3 | 58.3±7.5 | **29.9±6.3** | 48.7 | 87.7±11.0 | 69.7±7.6 | **31.3±6.5** | 55.0 | 103.8±15.5 | 77.3±8.8 | **33.8±6.5** | 56.2 |
| german1 | 80.0±10.6 | 75.9±6.4 | **34.5±5.6** | 54.5 | 149.3±71.8 | 93.2±11.3 | **34.9±5.7** | 62.5 | 211.1±141.7 | 100.0±0.0 | **35.3±5.8** | 64.7 |
| haberman1 | 35.9±5.4 | 27.0±0.0 | **13.8±2.7** | 48.8 | 55.1±9.4 | 27.0±0.0 | **12.5±2.8** | 53.8 | 87.1±16.2 | 27.0±0.0 | **12.7±2.8** | 53.0 |
| heart-s1 | 23.1±5.8 | 21.9±5.3 | **8.6±2.6** | 60.6 | 33.6±8.1 | 28.2±5.9 | **9.4±2.9** | 66.8 | 40.7±9.3 | 36.0±7.3 | **10.6±2.7** | 70.5 |
| ionosph1 | 7.6±2.4 | 6.7±2.4 | **0.3±0.5** | 95.5 | 9.1±3.2 | 8.3±3.2 | **0.3±0.5** | 96.4 | 11.7±4.9 | 9.9±4.8 | **0.3±0.5** | 97.0 |
| liver1 | 47.0±6.2 | 41.6±5.8 | **16.6±2.6** | 60.1 | 62.2±14.7 | 47.4±2.2 | **18.1±2.7** | 61.8 | 86.3±26.9 | 48.0±0.0 | **20.2±3.6** | 58.0 |
| sonar1 | 14.3±4.3 | 15.3±4.1 | **4.7±2.0** | 69.1 | 20.4±6.8 | 20.3±4.0 | **4.7±1.9** | 76.7 | 30.4±11.2 | 20.9±6.4 | **5.2±2.0** | 75.2 |
| spambase1 | 42.5±6.0 | 41.1±5.8 | **12.0±3.1** | 70.8 | 66.6±10.2 | 56.4±8.6 | **14.8±3.4** | 73.8 | 106.8±17.6 | 76.5±14.0 | **13.6±3.6** | 82.2 |
| spect1 | 12.0±2.7 | 11.5±3.2 | **4.8±1.8** | 58.3 | 13.3±3.8 | 13.0±3.2 | **5.3±2.0** | 59.2 | 15.5±6.2 | 13.4±3.6 | **6.2±1.9** | 53.7 |
| spectf1 | 13.2±2.9 | 13.2±2.9 | **4.9±2.6** | 63.2 | 13.2±2.9 | 13.2±2.9 | **5.5±2.6** | 58.7 | 13.2±2.9 | 13.2±2.9 | **5.0±2.6** | 62.2 |
| tic-tac1 | 4.7±1.3 | 4.7±1.3 | **1.1±1.4** | 76.8 | 4.7±1.3 | 4.7±1.3 | **1.5±1.5** | 67.6 | 4.7±1.3 | 4.7±1.3 | **1.4±1.4** | 71.1 |
| avg. | *33.9* | *31.1* | *12.5* | *63.1* | *59.3* | *36.8* | *13.7* | *64.0* | *90.1* | *41.5* | *14.3* | *64.6* |
| | | U(5,7) | | | | U(5,10) | | | | U(5,15) | | |
| breast-c1 | 106.0±57.2 | 28.0±0.0 | **11.9±2.5** | 57.6 | 130.0±74.1 | 28.0±0.0 | **11.9±2.5** | 57.6 | 170.0±102.3 | 28.0±0.0 | **11.9±2.5** | 57.6 |
| breast-w1 | 10.0±2.6 | 5.4±3.1 | **3.6±1.5** | 32.9 | 10.0±2.6 | 5.8±3.7 | **3.6±1.5** | 38.9 | 10.0±2.6 | 6.6±4.8 | **3.5±1.5** | 47.2 |
| credit-a1 | 113.8±78.1 | 62.9±8.6 | **19.1±4.4** | 69.6 | 131.2±100.1 | 66.9±10.7 | **19.1±4.3** | 71.4 | 160.3±136.8 | 74.0±11.6 | **19.4±4.4** | 73.8 |
| credit-g1 | 410.0±236.0 | 99.1±3.0 | **37.6±5.0** | 62.0 | 503.9±306.4 | 100.0±0.0 | **34.3±5.2** | 65.7 | 660.4±423.6 | 100.0±0.0 | **34.3±5.2** | 65.7 |
| diabetes1 | 103.8±15.5 | 77.3±8.8 | **37.7±6.7** | 51.2 | 116.0±19.6 | 82.3±8.8 | **38.5±6.7** | 53.3 | 136.3±26.8 | 89.0±0.0 | **40.4±6.3** | 54.6 |
| german1 | 211.1±141.7 | 100.0±0.0 | **35.5±5.8** | 64.5 | 248.2±183.6 | 100.0±0.0 | **35.7±5.7** | 64.3 | 310.1±253.6 | 100.0±0.0 | **35.7±5.7** | 64.3 |
| haberman1 | 87.1±16.2 | 27.0±0.0 | **12.8±2.8** | 52.6 | 106.3±20.3 | 27.0±0.0 | **12.9±2.8** | 52.2 | 138.3±27.3 | 27.0±0.0 | **13.0±3.0** | 51.9 |
| heart-s1 | 40.7±9.3 | 36.0±7.3 | **8.1±2.4** | 77.5 | 45.2±11.8 | 38.1±6.2 | **8.2±2.4** | 78.4 | 52.7±16.0 | 39.4±5.2 | **8.4±2.3** | 78.8 |
| ionosph1 | 11.7±4.9 | 9.9±4.8 | **0.3±0.5** | 97.0 | 13.3±6.0 | 10.1±5.9 | **0.3±0.5** | 97.0 | 15.8±8.0 | 17.5±9.2 | **0.3±0.5** | 98.3 |
| liver1 | 86.3±26.9 | 48.0±0.0 | **22.6±3.1** | 52.8 | 100.8±34.3 | 48.0±0.0 | **23.6±3.6** | 50.8 | 125.0±46.6 | 48.0±0.0 | **25.1±4.2** | 47.8 |
| sonar1 | 30.4±11.2 | 20.9±6.4 | **5.2±2.0** | 75.1 | 36.5±13.9 | 23.4±8.0 | **6.3±2.2** | 73.2 | 46.6±18.4 | 25.9±9.9 | **6.6±2.3** | 74.4 |
| spambase1 | 106.8±17.6 | 76.5±14.0 | **16.1±4.4** | 79.0 | 130.9±22.1 | 83.4±15.6 | **16.8±5.4** | 79.9 | 156.1±37.4 | 87.5±13.2 | **21.4±8.8** | 75.6 |
| spect1 | 15.5±6.2 | 13.4±3.6 | **5.8±2.2** | 56.7 | 16.8±7.7 | 15.3±3.7 | **5.9±2.1** | 61.7 | 19.0±10.2 | 18.5±2.8 | **6.2±2.6** | 66.7 |
| spectf1 | 13.2±2.9 | 13.2±2.9 | **5.0±2.6** | 62.2 | 13.2±2.9 | 13.2±2.9 | **5.0±2.6** | 62.2 | 13.2±2.9 | 13.2±2.9 | **5.0±2.6** | 62.2 |
| tic-tac1 | 4.7±1.3 | 4.7±1.3 | **1.4±1.4** | 71.1 | 4.7±1.3 | 4.7±1.3 | **1.4±1.4** | 71.1 | 4.7±1.3 | 4.7±1.3 | **1.4±1.4** | 71.1 |
| avg. | *90.1* | *41.5* | *14.8* | *64.1* | *107.1* | *43.1* | *14.9* | *65.2* | *134.6* | *45.3* | *15.5* | *66.0* |

# Conclusion

LAMDA
Learning And Mining from DatA
http://lamda.nju.edu.cn

## Main contribution:

- ✓ The first study on learning with cost intervals
- ✓ The CISVM method

  http://lamda.nju.edu.cn/code_CISVM.ashx

## Future work:

- ✓ To improve the surrogate risk
- ✓ . . .

Thanks!