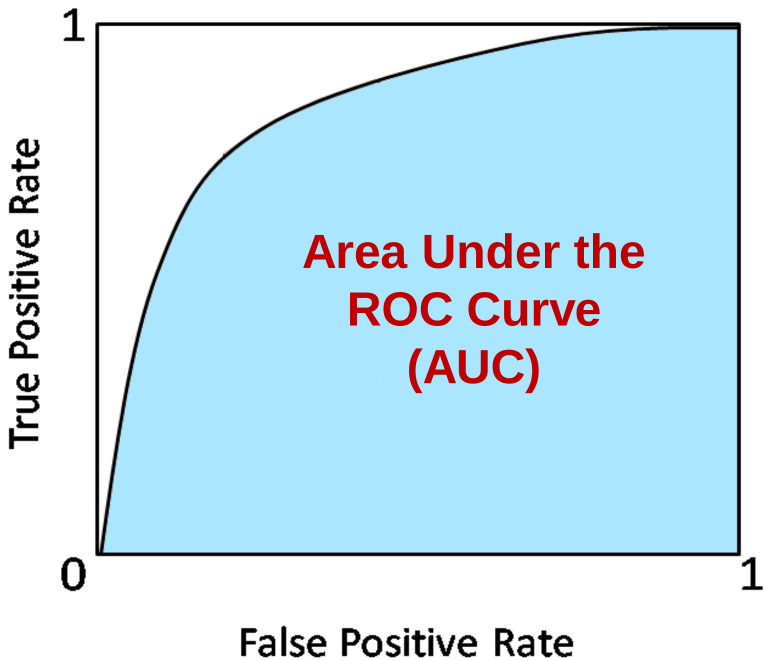
Support Vector Algorithms for Optimizing the Partial AUC

Harikrishna Narasimhan and Shivani Agarwal

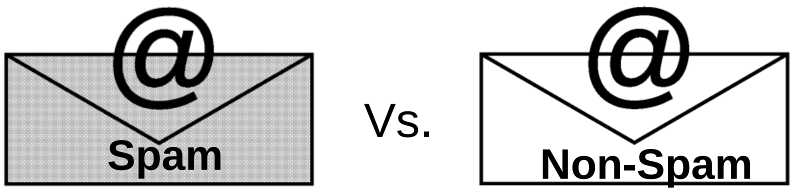
Department of Computer Science and Automation
Indian Institute of Science, Bangalore

Based on work in ICML 2013 and KDD 2013

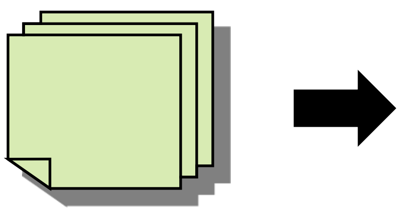
Receiver Operating Characteristic Curve



Binary Classification

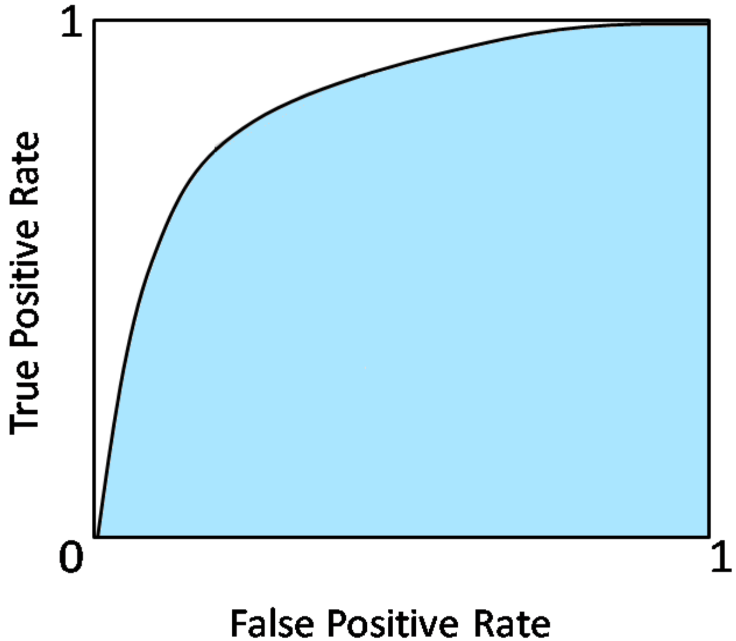


Bipartite Ranking

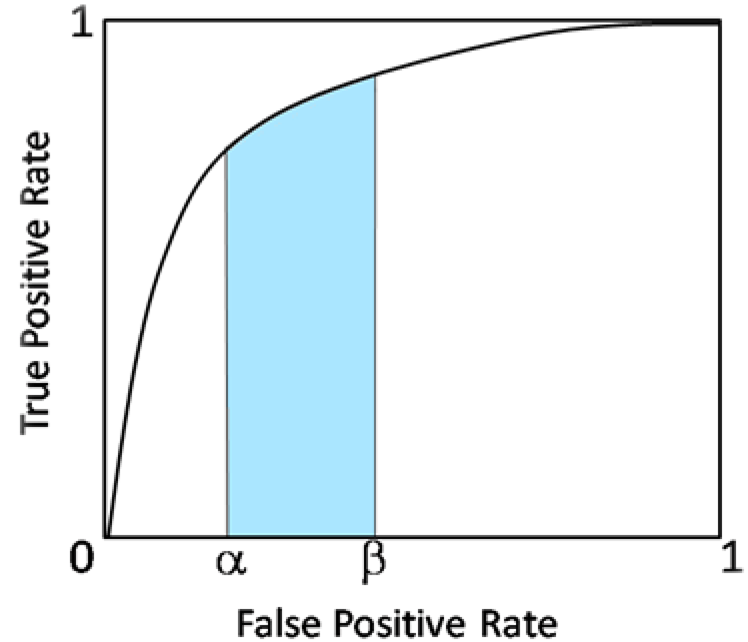


Ranking of documents

Partial AUC?



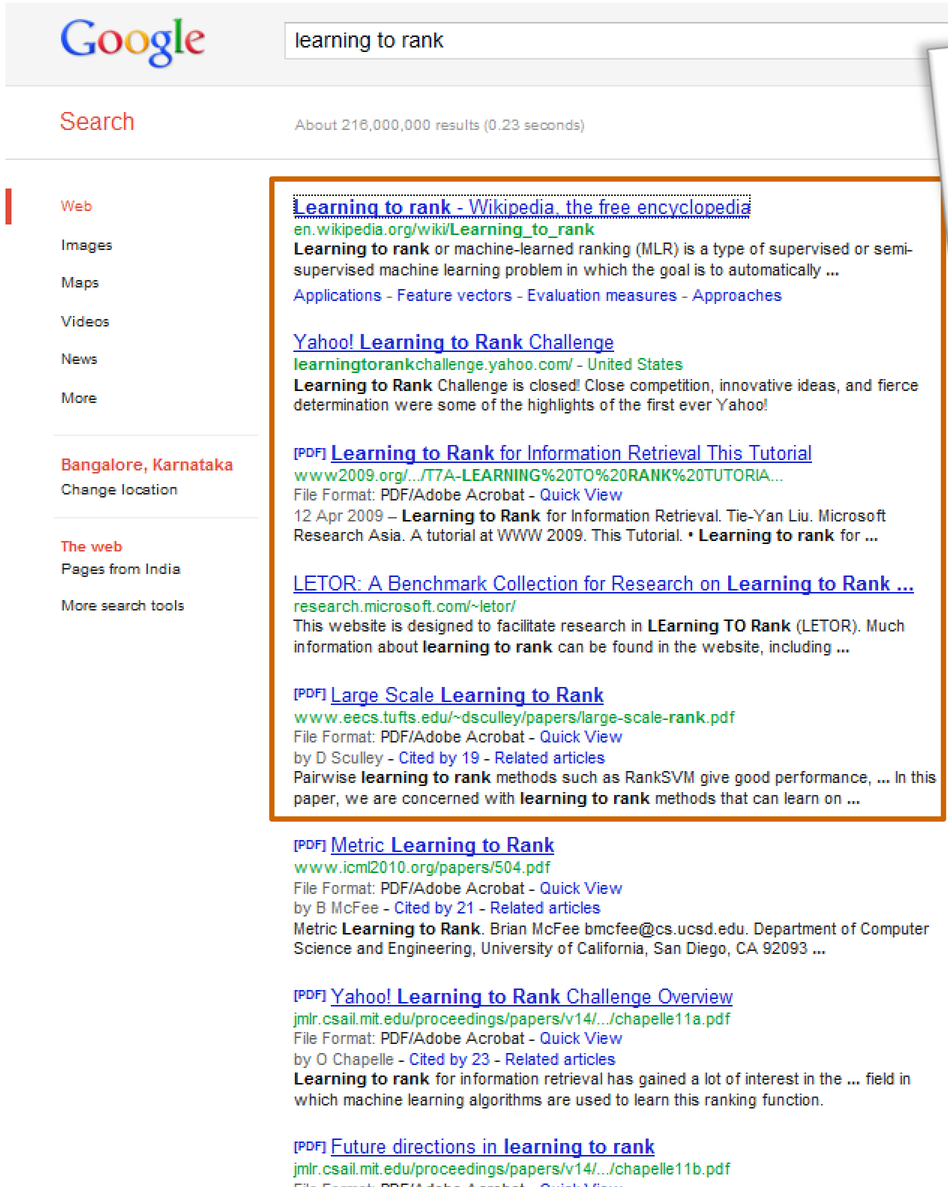
Vs



Full AUC

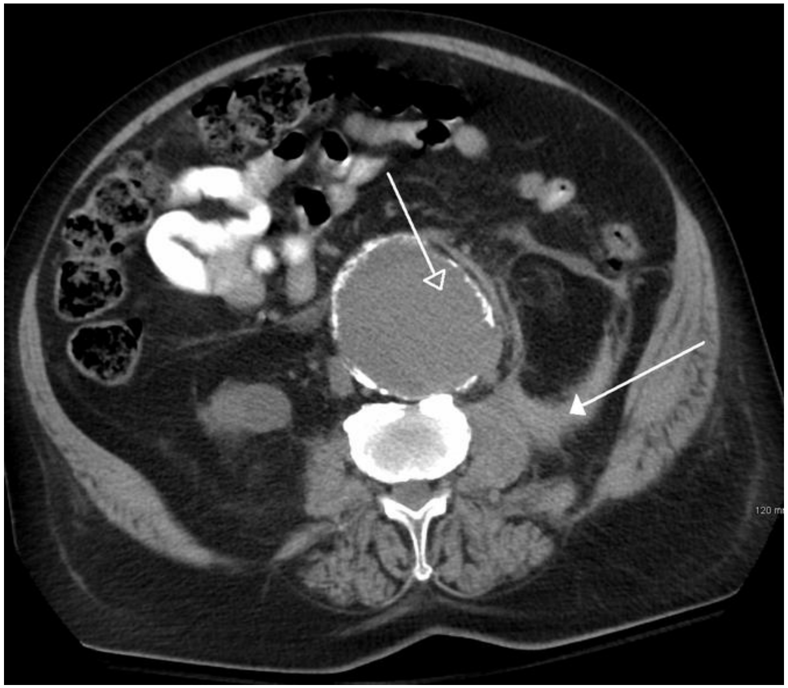
Partial AUC

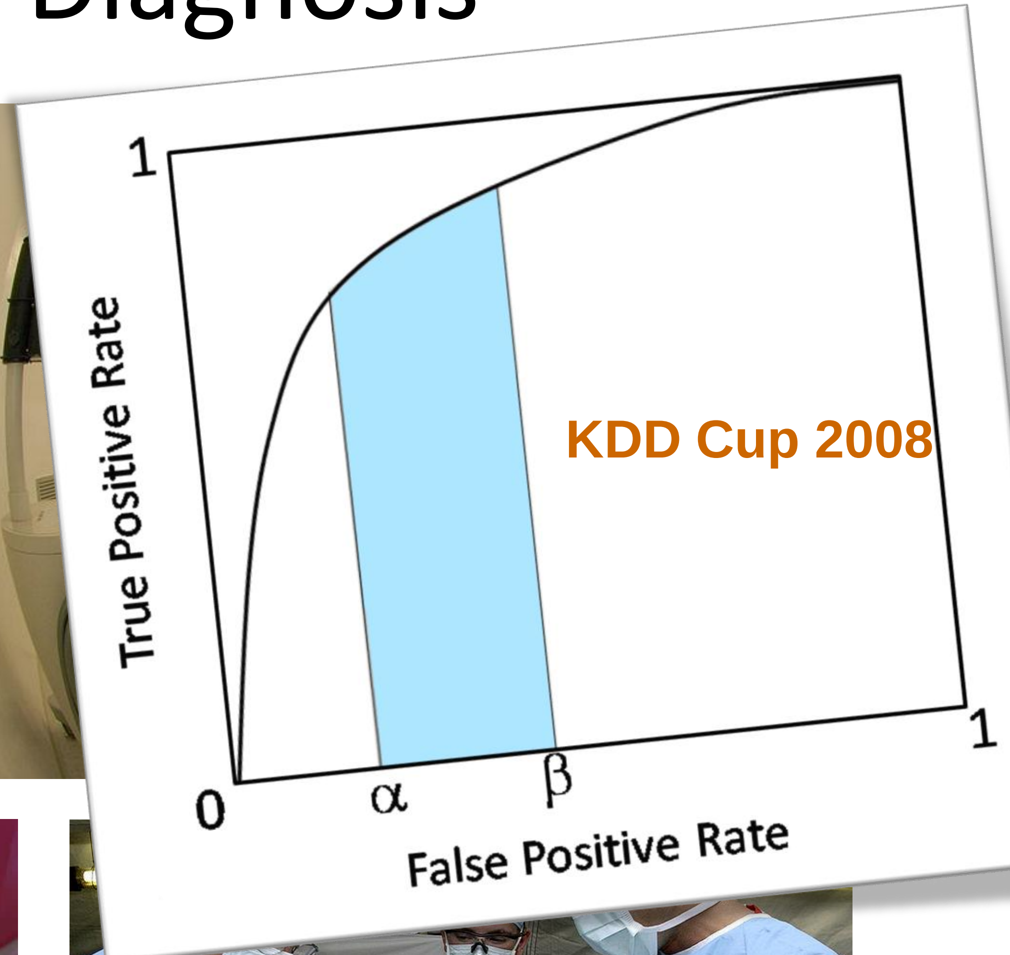
Ranking

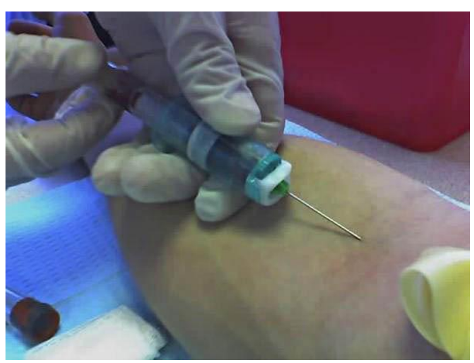


http://www.google.com/

Medical Diagnosis



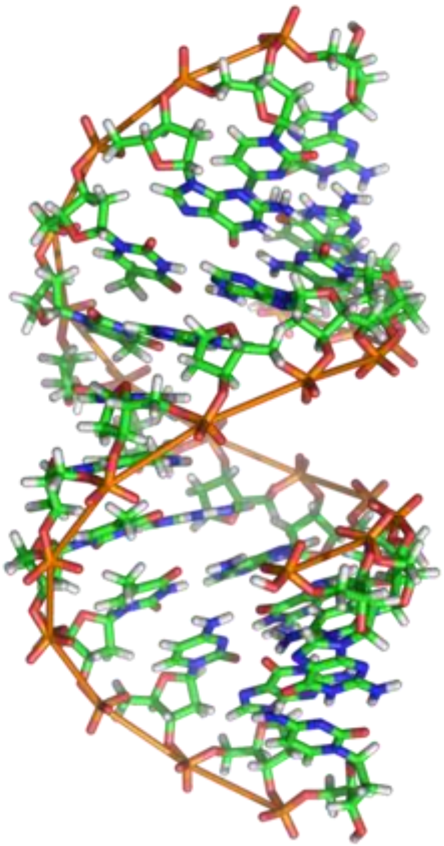


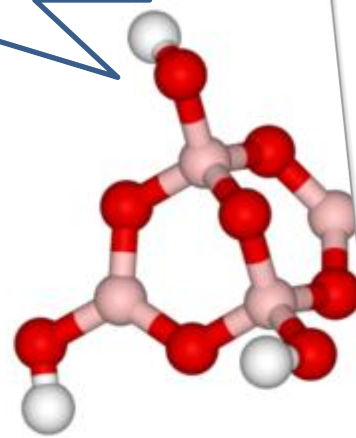


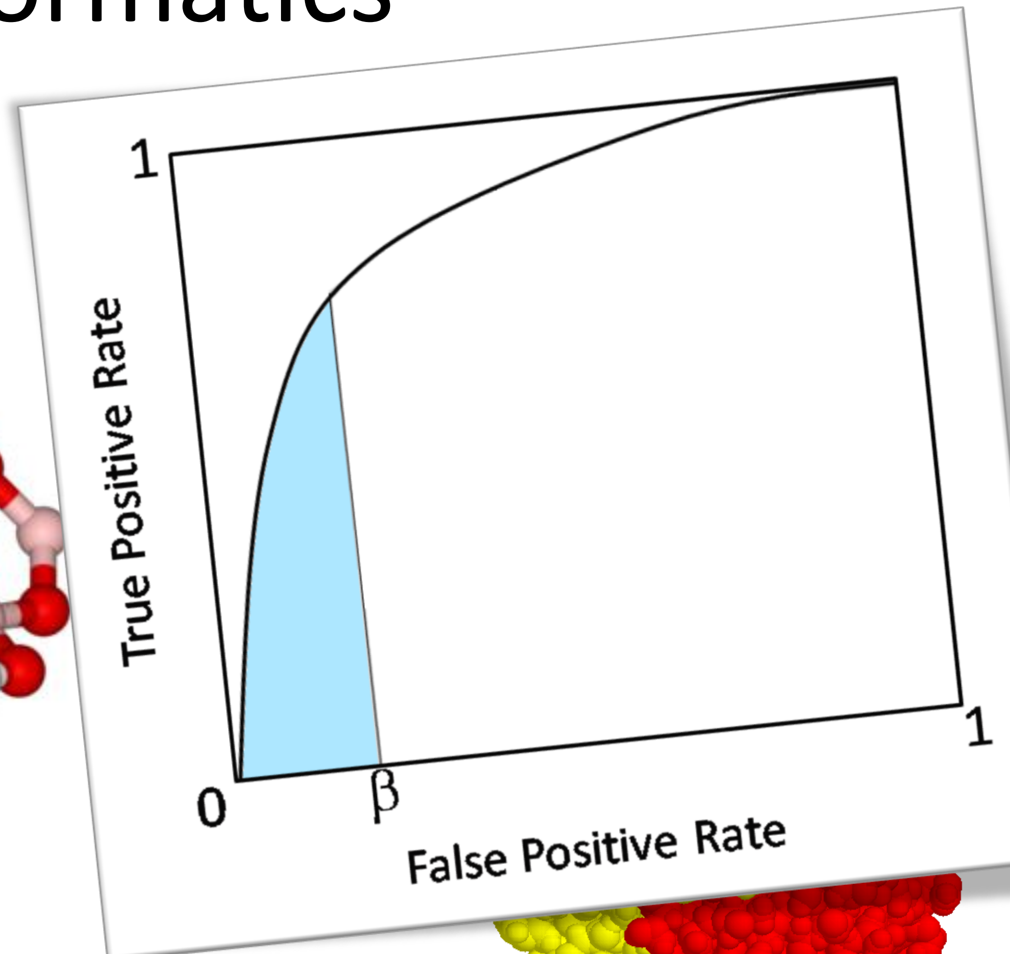
http://en.wikipedia.org/

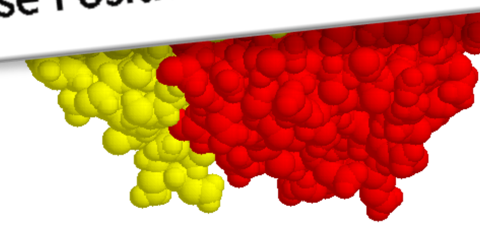
Bioinformatics

- Drug Discovery
- Gene Prioritization
- Protein Interaction Prediction
- ……









http://en.wikipedia.org/wiki
http://commons.wikimedia.org/
http://www.google.com/imghp

Partial Area Under the ROC Curve is critical to many applications

Partial AUC Optimization

- Asymmetric SVM:
  - Wu, S.-H., Lin, K.-P., Chen, C.-M., and Chen, M.-S. Asymmetric support vector machines: low false-positive learning under the user tolerance. In KDD, 2008.
- Boosting style algorithm:
  - Komori, O. and Eguchi, S. A boosting method for maximizing the partial area under the ROC curve. *BMC Bioinformatics*, 11:314, 2010.
  - Takenouchi, T., Komori, O., and Eguchi, S. An extension of the receiver operating characteristic curve and AUC-optimal classification. Neural Computation, 24, (10):2789–2824, 2012.
- Several heuristic approaches:
  - Pepe, M. S. and Thompson, M. L. Combining diagnostic test results to increase accuracy. *Biostatistics*, 1(2):123–140, 2000.
  - Ricamato, M. T. and Tortorella, F. Partial AUC maximization in a linear combination of dichotomizers. *Pattern Recognition*, 44(10-11):2669–2677, 2011.

Partial AUC Optimization

- Many of the existing approaches are either heuristic or solve special cases of the problem.
- **Our contribution**: New support vector methods for optimizing the general partial AUC measure.
- Based on Joachims' Structural SVM approach for optimizing full AUC, but leads to a trickier inner combinatorial optimization problem.
  - Joachims, T. A Support Vector Method for Multivariate Performance Measures. ICML, 2005.
  - Joachims, T. Training linear SVMs in linear time. KDD, 2006.
- Improvements over baselines on several real-world applications

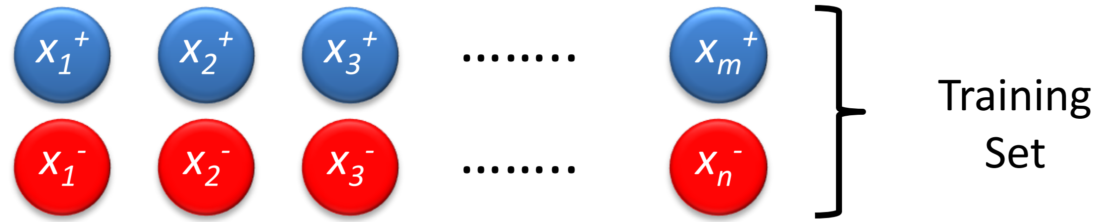
Outline

- Problem Setup
- **First cut:** Structural SVM Approach for Optimizing Partial AUC
- **Better Formulation:** Tighter Upper Bound on the Partial AUC Loss
- Experiments

Receiver Operating Characteristic Curve

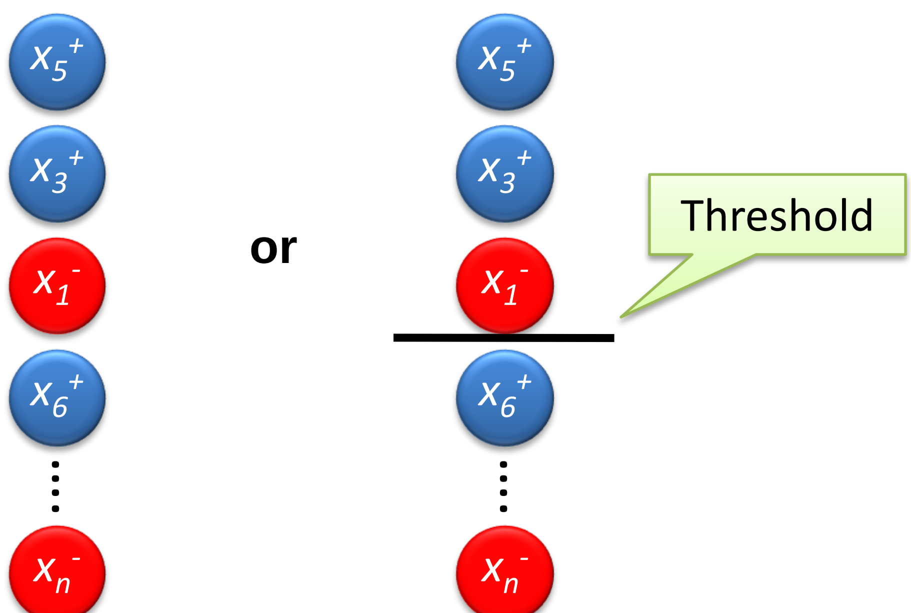
Positive Instances

Negative Instances

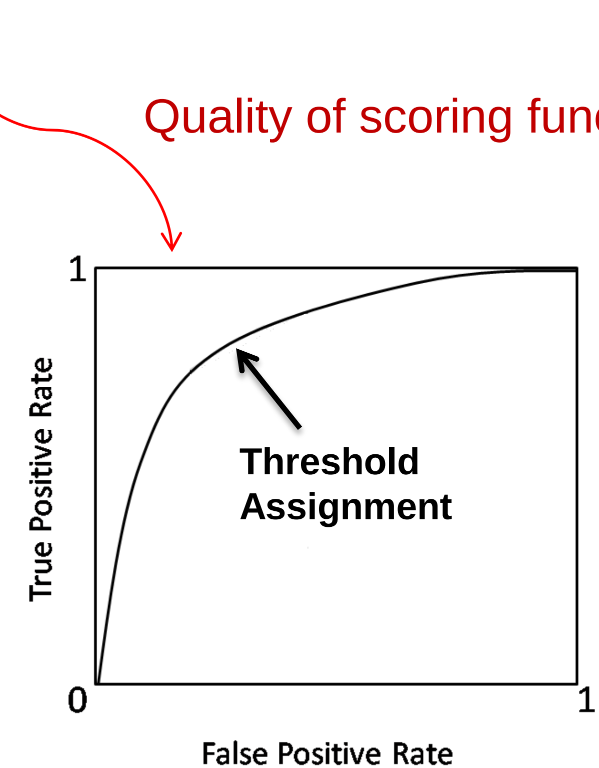


GOAL? Learn a scoring function $f : X \rightarrow \mathbb{R}$

Rank objects **Build a classifier**

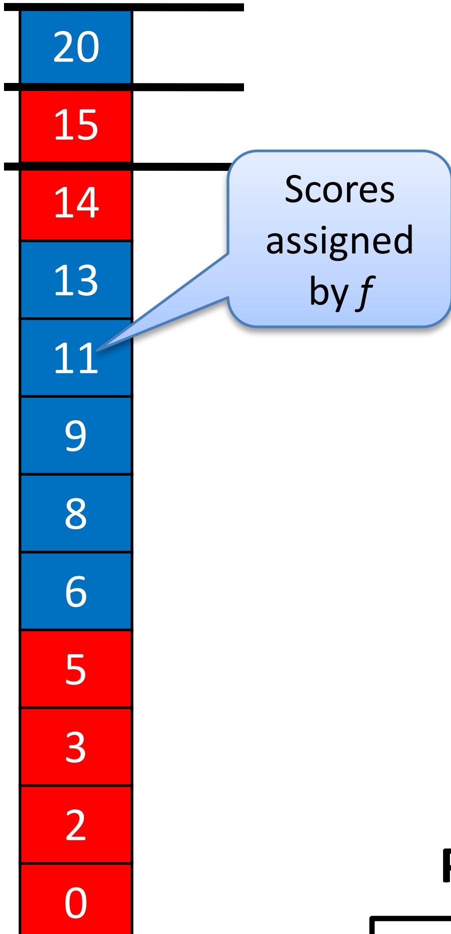


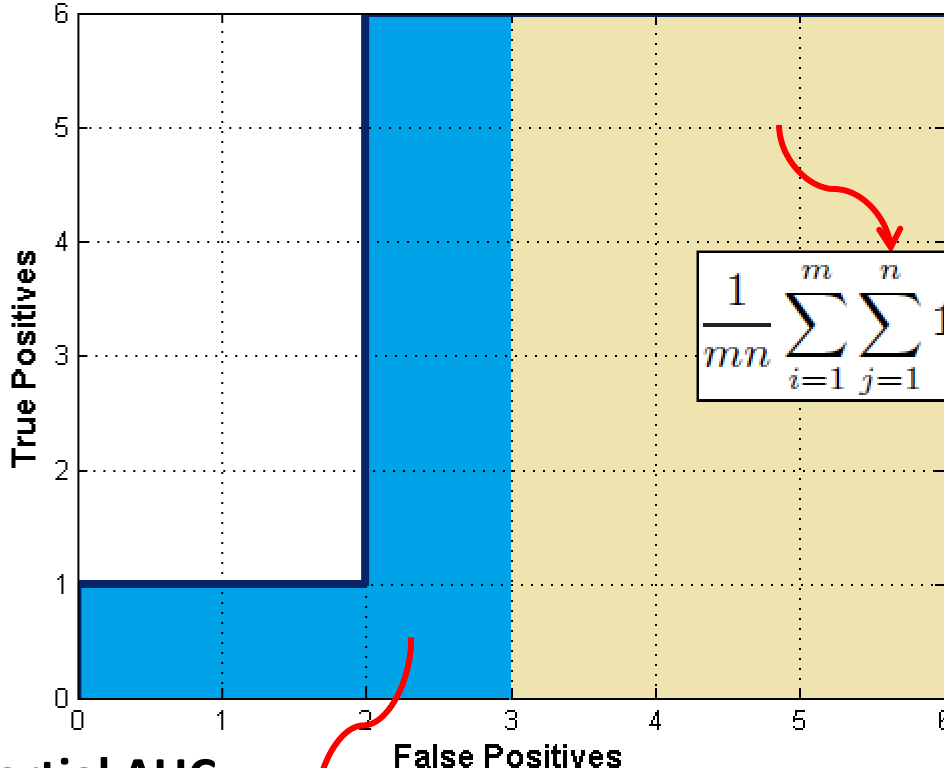
Quality of scoring function?



ROC Curve

Receiver Operating Characteristic Curve





Area Under the Curve (AUC)

$$\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{1}\left(f(x_i^+) > f(x_j^-)\right)$$

Pair-wise accuracy

Partial AUC

$$\frac{1}{mn(\beta-\alpha)}\sum_{i=1}^{m}\sum_{j=j_\alpha+1}^{j_\beta}\mathbf{1}\left(f(x_i^+) > f(x_{(j)}^-)\right)$$

Partial AUC Optimization

Minimize: $\underbrace{1 - \widehat{\text{pAUC}}_f(\alpha, \beta)}$

Discrete and
Non-differentiable

$$\underbrace{\text{Convex Upper Bound on ``}1 - \widehat{\text{pAUC}}_f(\alpha, \beta)\text{''} + \text{Regularizer}}_{\text{Structural SVM Based Approach}}$$

- Extends Joachims' approach for full AUC optimization, but leads to a trickier combinatorial optimization step.
- Efficient solver with the same/lesser time complexity compared to that for full AUC.

Outline

- Problem Setup
- **First cut:** Structural SVM Approach for Optimizing Partial AUC
- **Better Formulation:** Tighter Upper Bound on the Partial AUC Loss
- Experiments

Structural SVM Based Approach

Ordering of $\{x_1, x_2, \ldots, x_s\}$

$$\pi: \quad \begin{array}{c|c} & n \\ \hline m & \begin{array}{|c|c|c|c|c|} \hline 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 & 1 \\ \hline 1 & 1 & 0 & 0 & 1 \\ \hline \end{array} \end{array}$$

compared with

$$\pi^*: \quad \begin{array}{|c|c|c|c|c|} \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline \end{array}$$

IDEAL

Upper Bound on (1 – pAUC)

Regularizer

$$\min_{w,\xi \geq 0} \boxed{\frac{1}{2}||w||^2} + C\boxed{\xi}$$

s.t.

$$\boxed{\forall \pi \in \Pi_{m,n}} : \quad w^\top \left(\phi(S,\pi^*) - \phi(S,\pi)\right) \geq \Delta_{\mathrm{pAUC}(\alpha,\beta)}(\pi^*,\pi) - \xi$$

pAUC Loss

Exponential Number of Output Matrices!!

Cutting-plane Solver

Converges in constant number of iterations

$$\min_{w,\xi\geq 0} \frac{1}{2}||w||^2 + C\xi$$

s.t. $\forall \pi \in \mathcal{C}$:

$$\boxed{w^\top \left(\phi(S,\pi^*) - \phi(S,\pi)\right) \geq \Delta_{\mathrm{pAUC}(\alpha,\beta)}(\pi^*,\pi) - \xi}$$

Repeat:

1. Solve OP for a subset of constraints.
2. Add the most violated constraint.

Break down!

$$\underset{\pi}{\mathrm{argmax}} \ \Delta_{\mathrm{pAUC}(\alpha,\beta)}(\pi^*,\pi) + w^\top \left(\phi(S,\pi^*) - \phi(S,\pi)\right)$$

Full AUC

0	1	0	1	0
1	1	0	0	0
1	1	0	0	1
1	1	0	0	1

Partial AUC

0	1	0	1	0
1	1	0	0	0
1	1	0	0	1
1	1	0	0	1

?

Trickier Optimization Problem

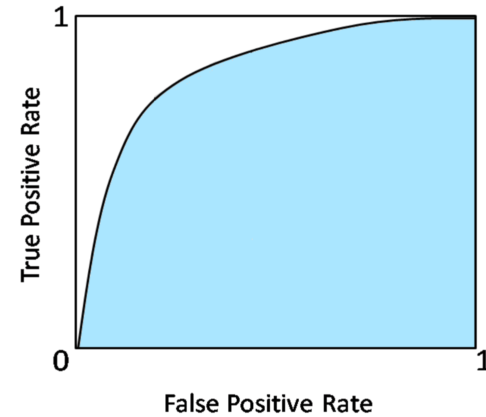
Full AUC

All Pairs

$$\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbb{1}\left(f(x_i^+) > f(x_j^-)\right)$$

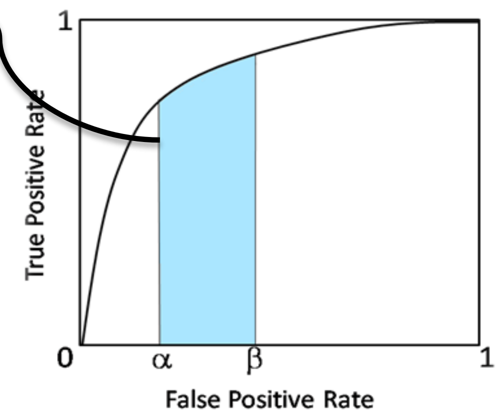
0	1	0	1	0
1	1	0	0	0
1	1	0	0	1
1	1	0	0	1

➡ $\sum$ ➡ **AUC**



Partial AUC

$$\frac{1}{mn(\beta-\alpha)}\sum_{i=1}^{m}\sum_{j=j_\alpha+1}^{j_\beta}\mathbb{1}\left(f(x_i^+) > f(x_{(j)}^-)\right)$$



Subset of negative instances in the FPR range [α, β] – **changes with ordering**

Trickier Optimization Problem

Full AUC

All Pairs

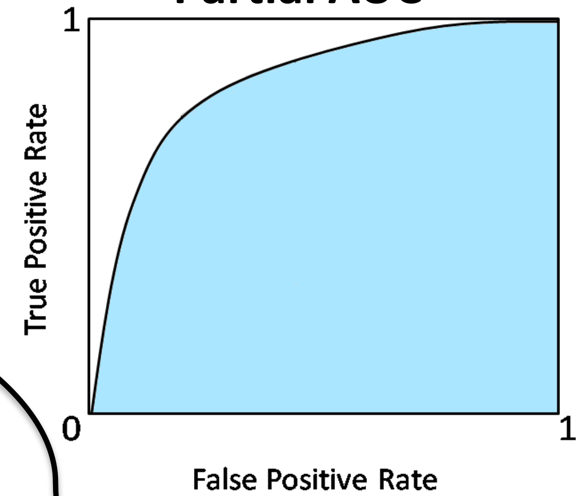
$$\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbb{1}\left(f(x_i^+) > f(x_j^-)\right)$$

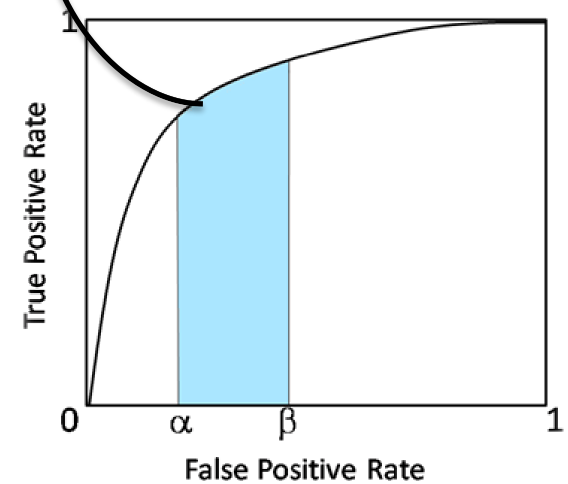
Partial AUC

$$\frac{1}{mn(\beta-\alpha)}\sum_{i=1}^{m}\sum_{j=j_\alpha+1}^{j_\beta}\mathbb{1}\left(f(x_i^+) > f(x_{(j)}^-)\right)$$

Subset of negative instances in the FPR range [α, β] – **changes with ordering**

Partial AUC





Can be implemented in **O((m+n) log (m+n))** time complexity

Inputs: $S = (S_+, S_-)$, α, β, w

For $i = 1, \ldots, m$ **do**

Optimize over $r_i \in \{0, \ldots, j_\alpha - 1\}$:

$$\pi^{(1)}_{i,(j)_w} = \begin{cases} \mathbf{1}(w^\top x^{\pm}_{i,(j)_w} \leq 0)\,, & j \in \{1, \ldots, j_\alpha - 1\} \\ 0\,, & j \in \{j_\alpha, \ldots, n\} \end{cases}$$

Optimize over $r_i \in \{j_\alpha\}$:

$$\pi^{(2)}_{i,(j)_w} = \begin{cases} 1\,, & j \in \{1, \ldots, j_\alpha\} \\ 0\,, & j \in \{j_\alpha + 1, \ldots, n\} \end{cases}$$

Optimize over $r_i \in \{j_\alpha + 1, \ldots, n\}$:

$$\pi^{(3)}_{i,(j)_w} = \begin{cases} 1\,, & j \in \{1, \ldots, j_\alpha + 1\} \\ \mathbf{1}(w^\top x^{\pm}_{i,(j)_w} \leq 1)\,, & j \in \{j_\alpha + 2, \ldots, j_\beta\} \\ \mathbf{1}(w^\top x^{\pm}_{i,(j)_w} \leq n\beta - j_\beta)\,, & j = j_\beta + 1 \\ \mathbf{1}(w^\top x^{\pm}_{i,(j)_w} \leq 0)\,, & j \in \{j_\beta + 2, \ldots, n\} \end{cases}$$

$$\bar{k} = \underset{k \in \{1,2,3\}}{\text{argmax}} \left\{ \begin{array}{l} \text{term inside sum over } i \text{ in} \\ \text{Eq. (4) evaluated at } \pi^{(k)}_i \end{array} \right\}$$

$\bar{\pi}_i = \pi^{(\bar{k})}_i$

End For

Output: $\bar{\pi}$

Partial AUC

0	1	0	1	0
1	1	0	0	0
1	1	0	0	1
1	1	0	0	1

Optimize rows independently

0	1	0	1	0
1	1	0	0	0
1	1	0	0	1
1	1	0	0	1

H. Narasimhan and S. Agarwal. *A Structural SVM Based Approach for Optimizing Partial AUC*. ICML, 2013.

Outline

- Problem Setup
- **First cut:** Structural SVM Approach for Optimizing Partial AUC
- **Better Formulation:** Tighter Upper Bound on the Partial AUC Loss
- Experiments

Better Formulation

- Tighter upper bound on partial AUC loss
- Lesser time for finding most-violated constraint!
- Better guarantee on number of cutting-plane iterations!

$$\max_{z \in \mathcal{Z}_\beta} \sum_{x_j^- \in z} \sum_{i=1}^{m} \mathbf{1}\big(f(x_i^+) < f(x_j^-)\big)$$

Max over subsets of negative instances

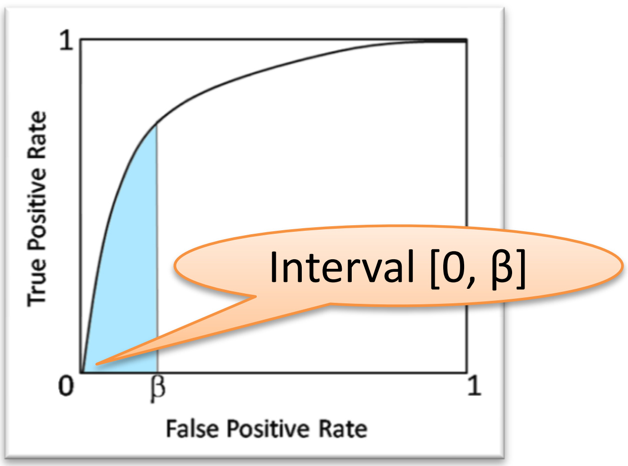
Truncated form of earlier objective

H. Narasimhan and S. Agarwal. *SVM_pAUC^tight: A New Support Vector Method for Optimizing Partial AUC Based on a Tight Convex Upper Bound*. KDD, 2013. *To appear*.

Outline

- Problem Setup
- **First cut:** Structural SVM Approach for Optimizing Partial AUC
- **Better Formulation:** Tighter Upper Bound on the Partial AUC Loss
- Experiments

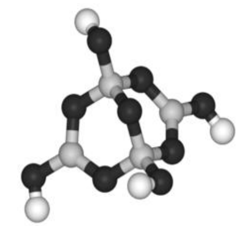
SVMpAUC^{struct} vs. Baseline Methods



Drug Discovery

50 active compounds / 2092 inactive compounds

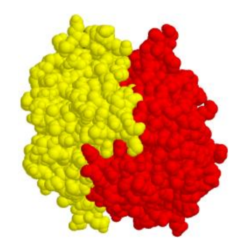
	pAUC(0, 0.1)
SVM$_{pAUC}$[0,0.1]	**65.25**
SVM$_{AUC}$	62.64 *
ASVM[0,0.1]	63.80
pAUCBoost[0,0.1]	43.89 *
Greedy-Heuristic[0,0.1]	8.33 *



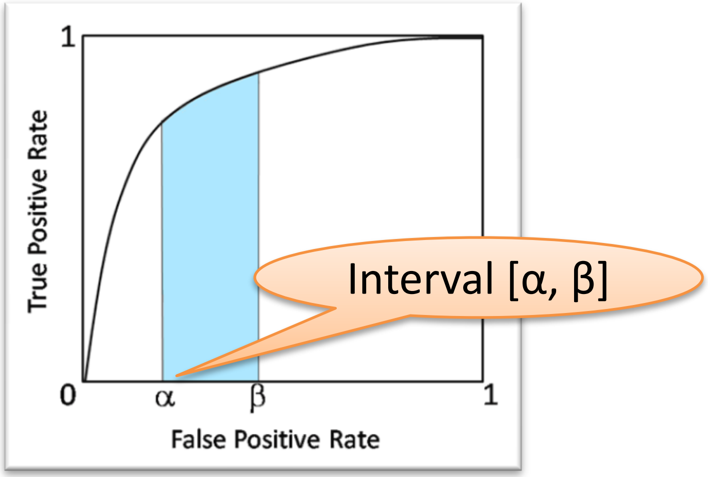
Protein-Protein Interaction Prediction

~3x10^3 interacting pairs / ~2x10^5 non-interacting pairs

	pAUC(0, 0.1)
SVM$_{pAUC}$[0,0.1]	**51.79**
SVM$_{AUC}$	39.72 *
ASVM[0,0.1]	44.51 *
pAUCBoost[0,0.1]	48.65 *
Greedy-Heuristic[0,0.1]	47.33 *



SVMpAUC^{struct} vs. Baseline Methods



KDD Cup 2008

Breast Cancer Detection

~600 malignant ROIs / ~10^5 benign ROIs

	pAUC$(0.2s, 0.3s)$
SVM$_{\text{pAUC}}[0.2s, 0.3s]$	51.44
SVM$_{\text{AUC}}$	50.50
pAUCBoost$[0.2s, 0.3s]$	48.06 *
Greedy-Heuristic$[0.2s, 0.3s]$	46.99 *

SVMpAUC^{tight} vs. SVMpAUC^{struct}

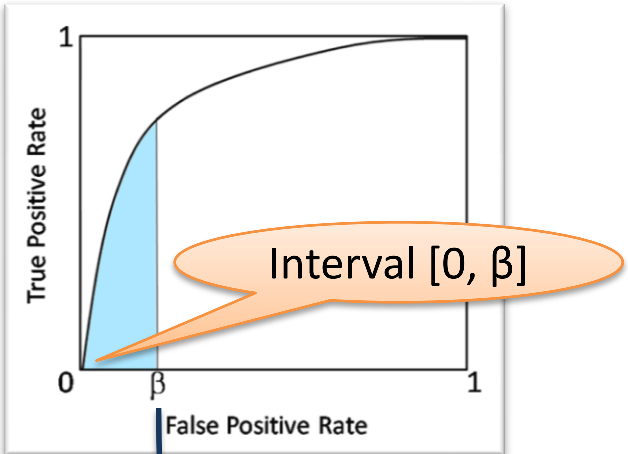
Partial AUC in [0, β]

	$\mathrm{SVM}_{\mathrm{pAUC}}^{\mathrm{tight}}[0, 0.1]$	$\mathrm{SVM}_{\mathrm{pAUC}}^{\mathrm{struct}}[0, 0.1]$	$\mathrm{SVM}_{\mathrm{AUC}}$
PPI	**52.95**	51.96 *	39.72 *
Cheminformatics	**65.30**	65.28	62.78
KDD Cup 2001	69.91	**70.12**	62.23 *
Leukemia	**30.44**	24.64 *	28.83
Ovarian Cancer	91.84	91.84	**92.17**

Partial AUC in [α, β]

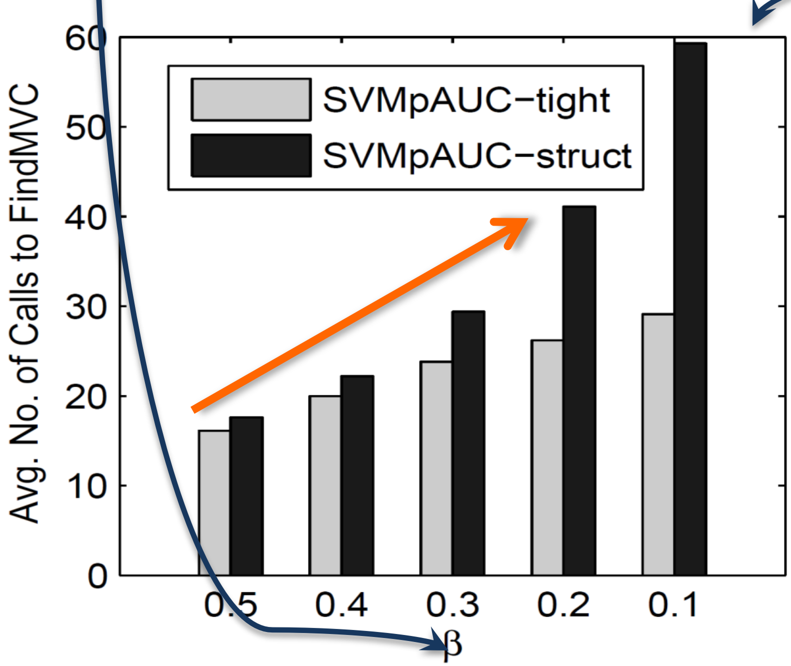
	$\mathrm{SVM}_{\mathrm{pAUC}}^{\mathrm{tight}}[0.2s, 0.3s]$	$\mathrm{SVM}_{\mathrm{pAUC}}^{\mathrm{struct}}[0.2s, 0.3s]$	$\mathrm{SVM}_{\mathrm{AUC}}$
KDD Cup 2008	**53.43**	51.89	50.66

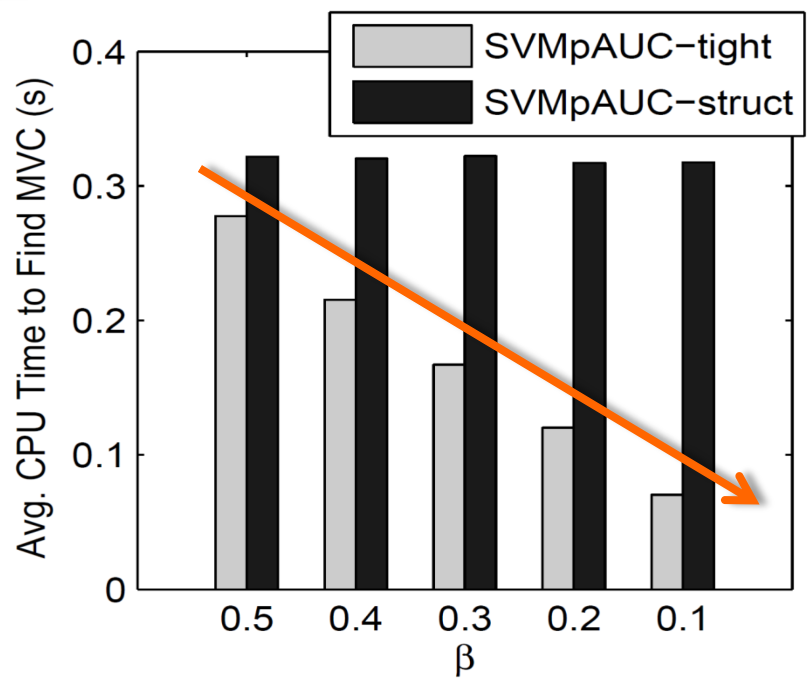
Run-time Analysis



Repeat:

1. Solve OP for a subset of constraints.
2. Add the most violated constraint.





Conclusions

- A new structural SVM based approach for optimizing partial AUC
- Efficient algorithm for solving the inner combinatorial optimization step
- Improved algorithm that optimizes a tighter upper bound on the partial AUC loss
- Experimental results confirm the effectiveness of our methods

Questions?