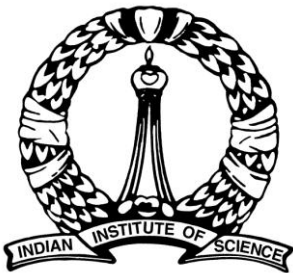


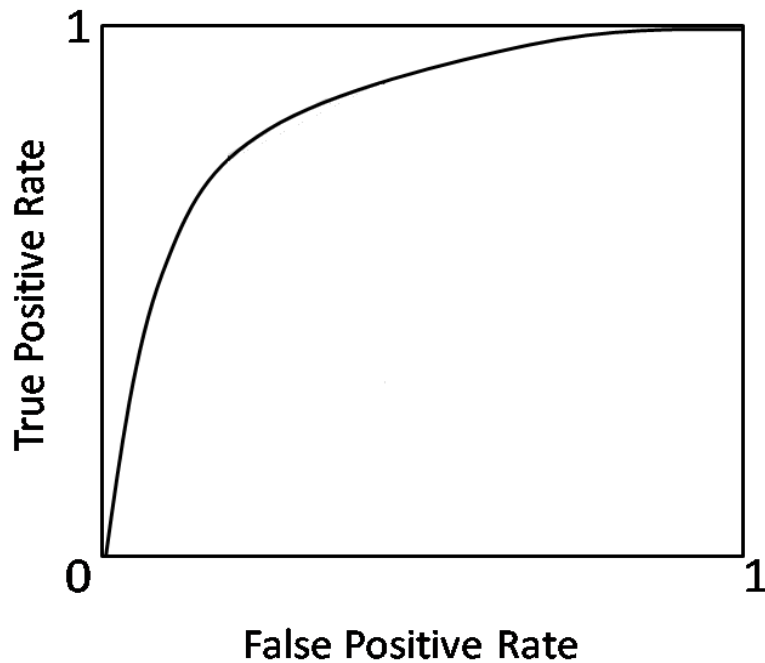
SVMpAUC-tight: A new algorithm for optimizing **partial AUC** based on a *tight convex upper bound*

Harikrishna Narasimhan and Shivani Agarwal

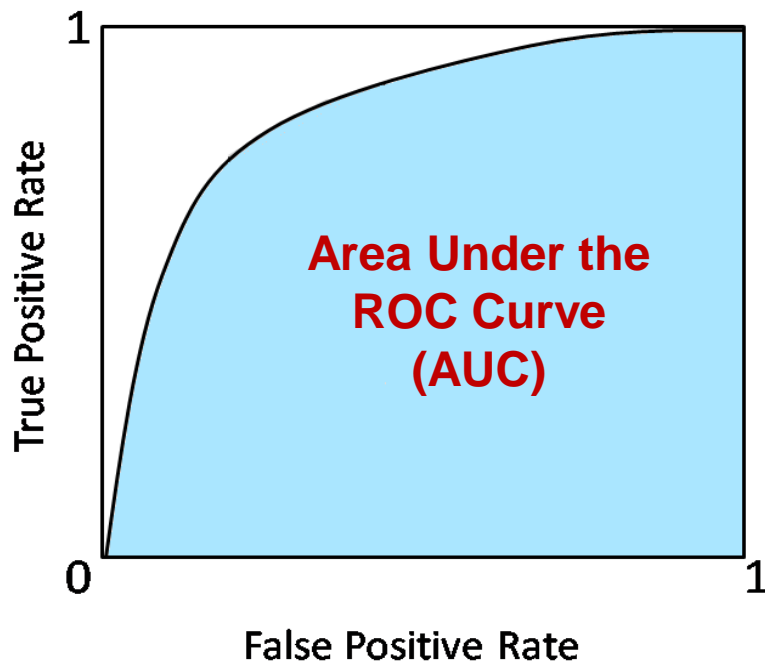


Department of Computer Science and Automation
Indian Institute of Science, Bangalore

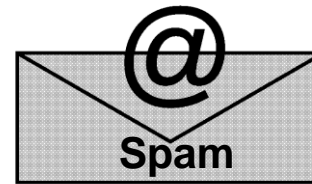
Receiver Operating Characteristic Curve



Receiver Operating Characteristic Curve



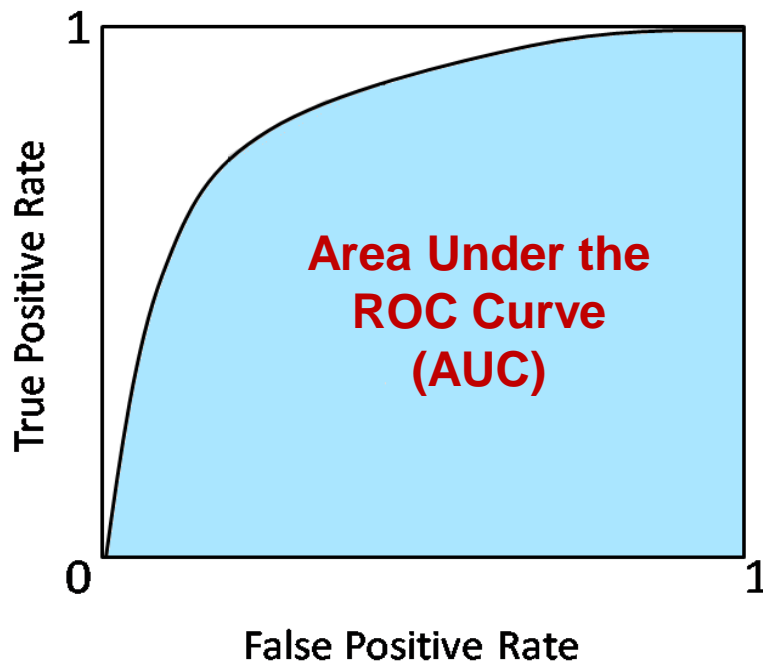
Binary Classification



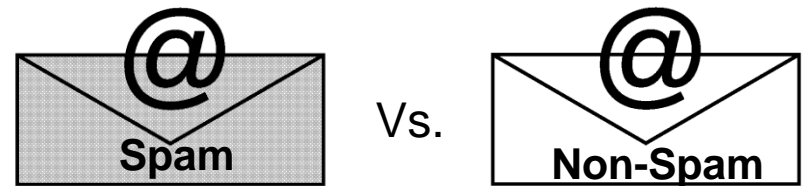
Vs.



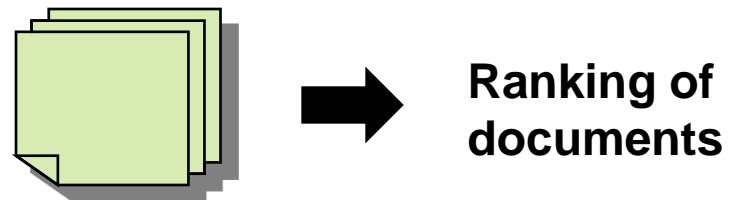
Receiver Operating Characteristic Curve



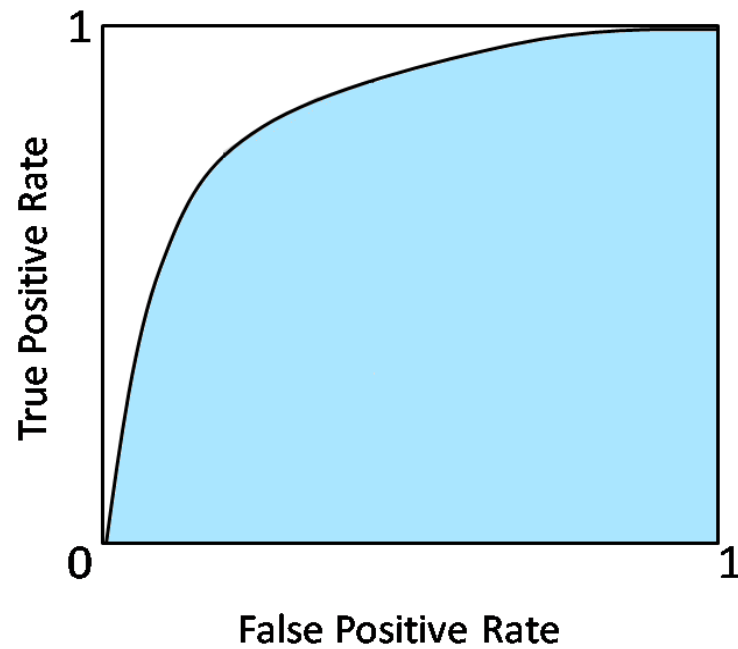
Binary Classification



Bipartite Ranking

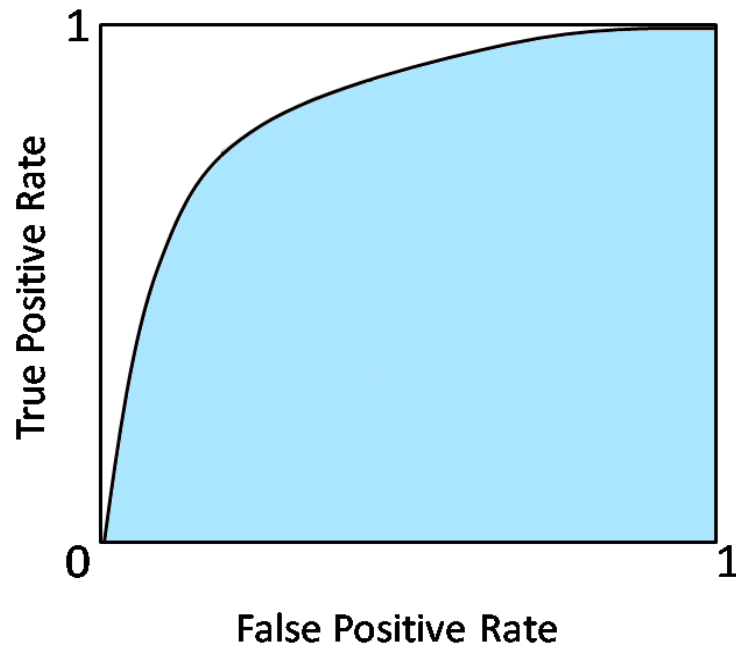


Partial AUC?



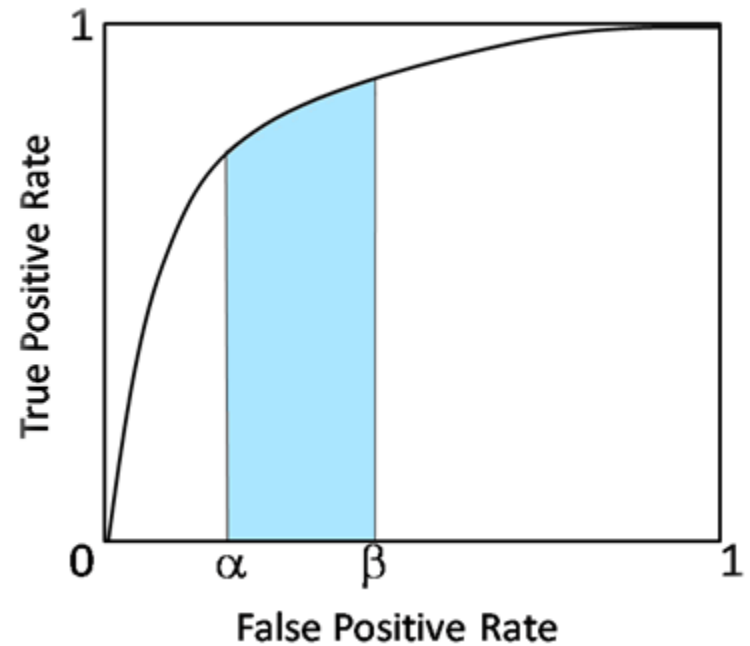
Full AUC

Partial AUC?



Full AUC

Vs



Partial AUC

Ranking



Search

About 216,000,000 results (0.23 seconds)

Web

Images

Maps

Videos

News

More

Bangalore, Karnataka

Change location

The web

Pages from India

More search tools

[Learning to rank - Wikipedia, the free encyclopedia](#)

en.wikipedia.org/wiki/Learning_to_rank

Learning to rank or machine-learned ranking (MLR) is a type of supervised or semi-supervised machine learning problem in which the goal is to automatically ...

[Applications](#) - [Feature vectors](#) - [Evaluation measures](#) - [Approaches](#)

[Yahoo! Learning to Rank Challenge](#)

learningtorankchallenge.yahoo.com/ - United States

Learning to Rank Challenge is closed! Close competition, innovative ideas, and fierce determination were some of the highlights of the first ever Yahoo!

[\[PDF\] Learning to Rank for Information Retrieval This Tutorial](#)

www.2009.org/.../T7A-LEARNING%20TO%20RANK%20TUTORIA...

File Format: PDF/Adobe Acrobat - [Quick View](#)

12 Apr 2009 - **Learning to Rank** for Information Retrieval. Tie-Yan Liu. Microsoft Research Asia. A tutorial at WWW 2009. This Tutorial. • **Learning to rank** for ...

[LETOR: A Benchmark Collection for Research on Learning to Rank ...](#)

research.microsoft.com/~letor/

This website is designed to facilitate research in **Learning TO Rank** (LETOR). Much information about **learning to rank** can be found in the website, including ...

[\[PDF\] Large Scale Learning to Rank](#)

www.eecs.tufts.edu/~dsculley/papers/large-scale-rank.pdf

File Format: PDF/Adobe Acrobat - [Quick View](#)

by D Sculley - [Cited by 19](#) - [Related articles](#)

Pairwise **learning to rank** methods such as RankSVM give good performance, ... In this paper, we are concerned with **learning to rank** methods that can learn on ...

[\[PDF\] Metric Learning to Rank](#)

www.icml2010.org/papers/504.pdf

File Format: PDF/Adobe Acrobat - [Quick View](#)

by B McFee - [Cited by 21](#) - [Related articles](#)

Metric Learning to Rank. Brian McFee bmcfee@cs.ucsd.edu. Department of Computer Science and Engineering, University of California, San Diego, CA 92093 ...

[\[PDF\] Yahoo! Learning to Rank Challenge Overview](#)

jmlr.csail.mit.edu/proceedings/papers/v14/.../chapelle11a.pdf

File Format: PDF/Adobe Acrobat - [Quick View](#)

by O Chapelle - [Cited by 23](#) - [Related articles](#)

Learning to rank for information retrieval has gained a lot of interest in the ... field in which machine learning algorithms are used to learn this ranking function.

[\[PDF\] Future directions in learning to rank](#)

jmlr.csail.mit.edu/proceedings/papers/v14/.../chapelle11b.pdf

File Format: PDF/Adobe Acrobat - [Quick View](#)

<http://www.google.com/>

Ranking



learning to rank

Search

About 216,000,000 results (0.23 seconds)

Web

Images

Maps

Videos

News

More

Bangalore, Karnataka

Change location

The web

Pages from India

More search tools

[Learning to rank - Wikipedia, the free encyclopedia](#)

en.wikipedia.org/wiki/Learning_to_rank

Learning to rank or machine-learned ranking (MLR) is a type of supervised or semi-supervised machine learning problem in which the goal is to automatically ...

[Applications](#) - [Feature vectors](#) - [Evaluation measures](#) - [Approaches](#)

[Yahoo! Learning to Rank Challenge](#)

learningtorankchallenge.yahoo.com/ - United States

Learning to Rank Challenge is closed! Close competition, innovative ideas, and fierce determination were some of the highlights of the first ever Yahoo!

[\[PDF\] Learning to Rank for Information Retrieval This Tutorial](#)

www2009.org/.../T7A-LEARNING%20TO%20RANK%20TUTORIA...

File Format: PDF/Adobe Acrobat - [Quick View](#)

12 Apr 2009 - **Learning to Rank** for Information Retrieval. Tie-Yan Liu. Microsoft Research Asia. A tutorial at WWW 2009. This Tutorial. • **Learning to rank** for ...

[LETOR: A Benchmark Collection for Research on Learning to Rank ...](#)

research.microsoft.com/~letor/

This website is designed to facilitate research in **Learning TO Rank** (LETOR). Much information about **learning to rank** can be found in the website, including ...

[\[PDF\] Large Scale Learning to Rank](#)

www.eecs.tufts.edu/~dsculley/papers/large-scale-rank.pdf

File Format: PDF/Adobe Acrobat - [Quick View](#)

by D Sculley - [Cited by 19](#) - [Related articles](#)

Pairwise **learning to rank** methods such as RankSVM give good performance, ... In this paper, we are concerned with **learning to rank** methods that can learn on ...

[\[PDF\] Metric Learning to Rank](#)

www.icml2010.org/papers/504.pdf

File Format: PDF/Adobe Acrobat - [Quick View](#)

by B McFee - [Cited by 21](#) - [Related articles](#)

Metric Learning to Rank. Brian McFee bmcfee@cs.ucsd.edu. Department of Computer Science and Engineering, University of California, San Diego, CA 92093 ...

[\[PDF\] Yahoo! Learning to Rank Challenge Overview](#)

jmlr.csail.mit.edu/proceedings/papers/v14/.../chapelle11a.pdf

File Format: PDF/Adobe Acrobat - [Quick View](#)

by O Chapelle - [Cited by 23](#) - [Related articles](#)

Learning to rank for information retrieval has gained a lot of interest in the ... field in which machine learning algorithms are used to learn this ranking function.

[\[PDF\] Future directions in learning to rank](#)

jmlr.csail.mit.edu/proceedings/papers/v14/.../chapelle11b.pdf

File Format: PDF/Adobe Acrobat - [Quick View](#)

True Positive Rate

1

0

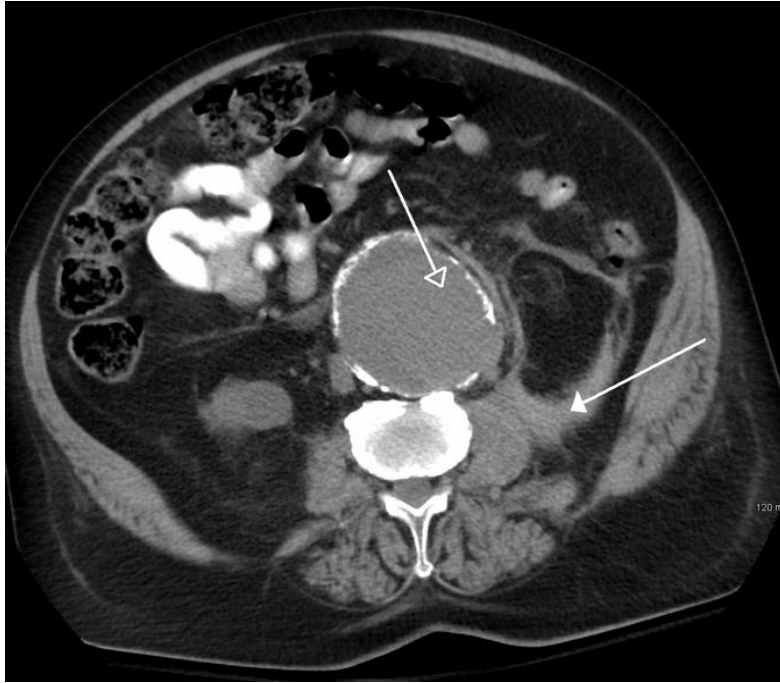
β

False Positive Rate

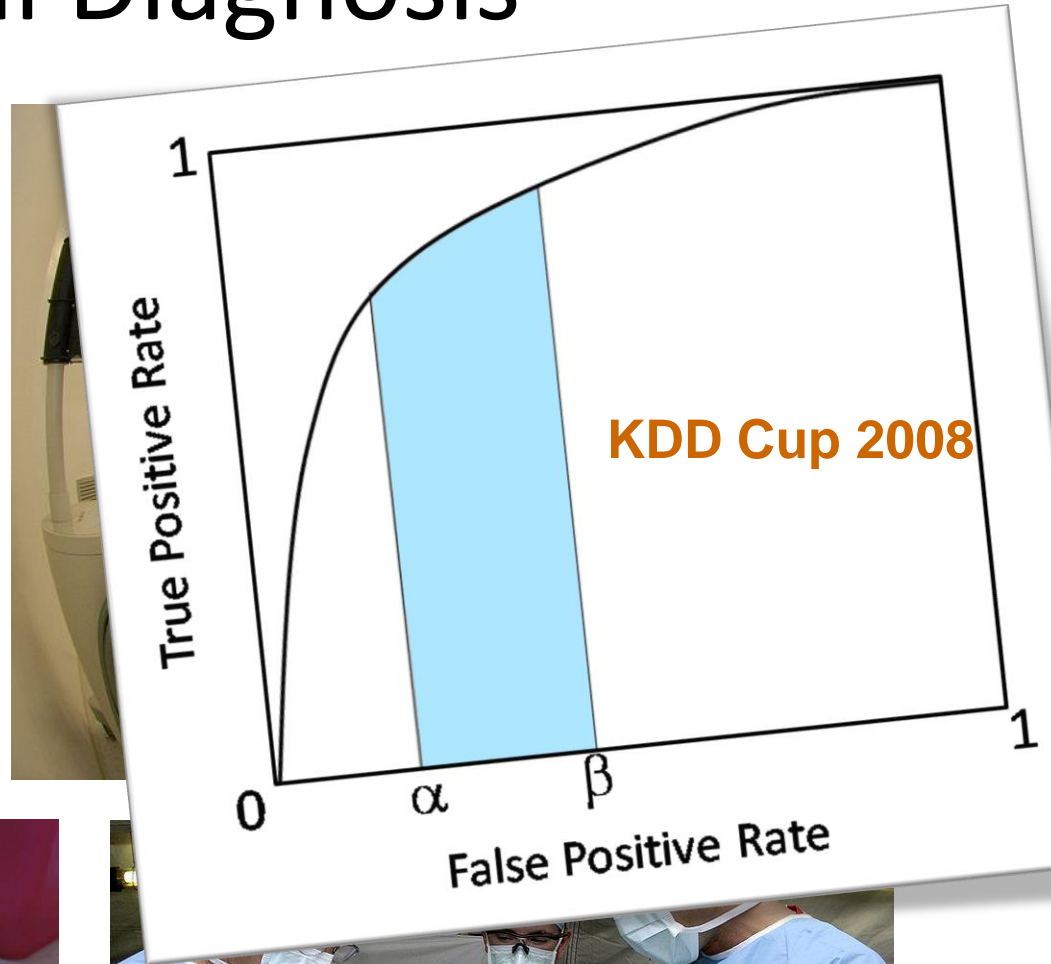
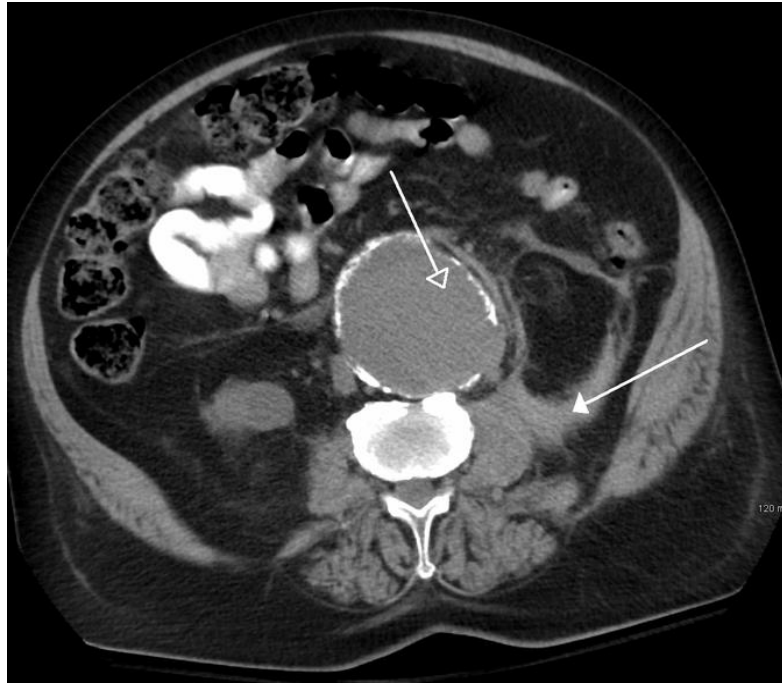
1

<http://www.google.com/>

Medical Diagnosis

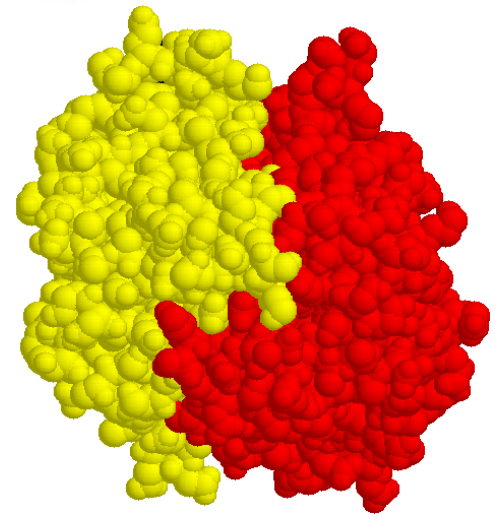
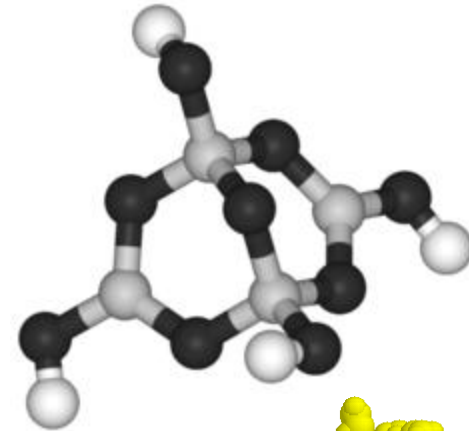
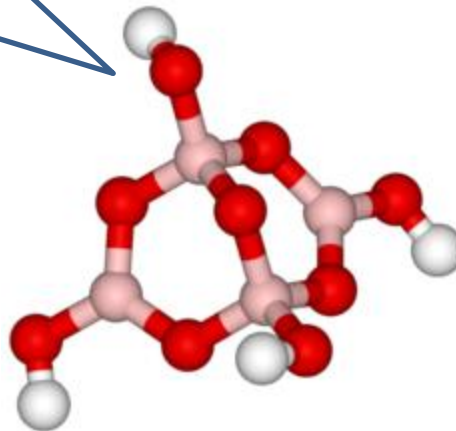
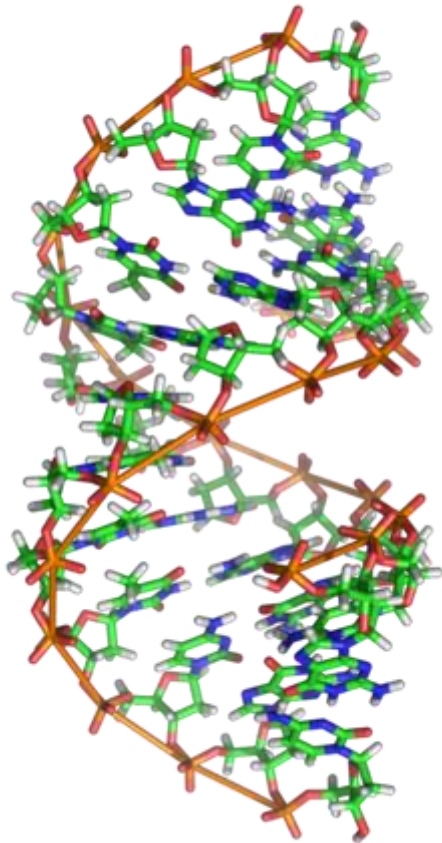


Medical Diagnosis



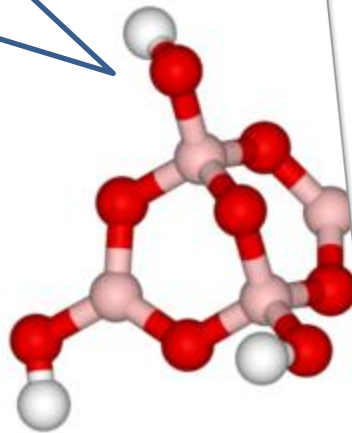
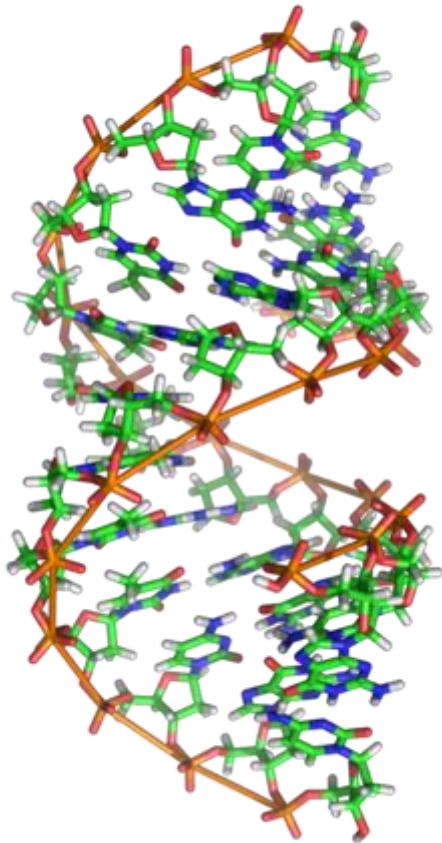
Bioinformatics

- Drug Discovery
- Gene Prioritization
- Protein Interaction Prediction
-



Bioinformatics

- Drug Discovery
- Gene Prioritization
- Protein Interaction Prediction
-



True Positive Rate

1

0

β

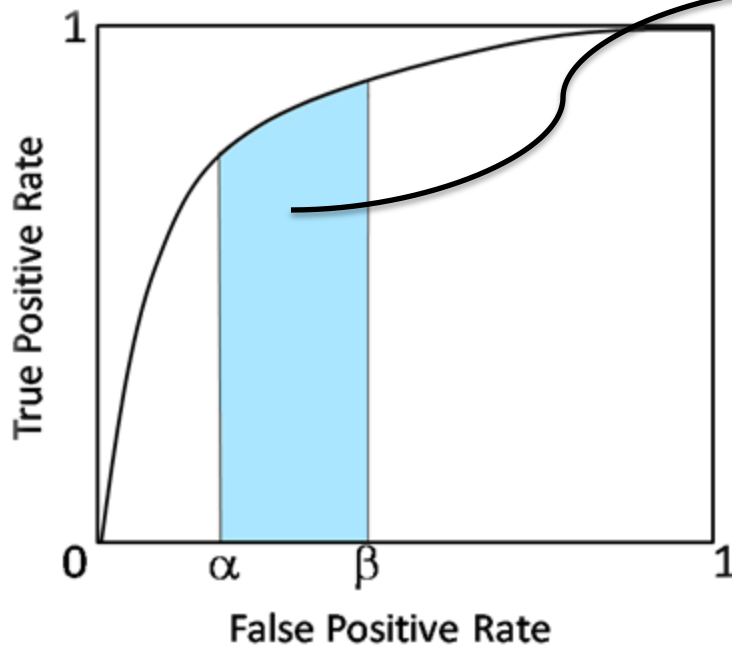
False Positive Rate

1

Partial Area Under the ROC Curve is critical
to many applications

SVMpAUC (ICML 2013)

Narasimhan, H. and Agarwal, S. “A *structural SVM* based approach for optimizing partial AUC”, ICML 2013.



SVMpAUC

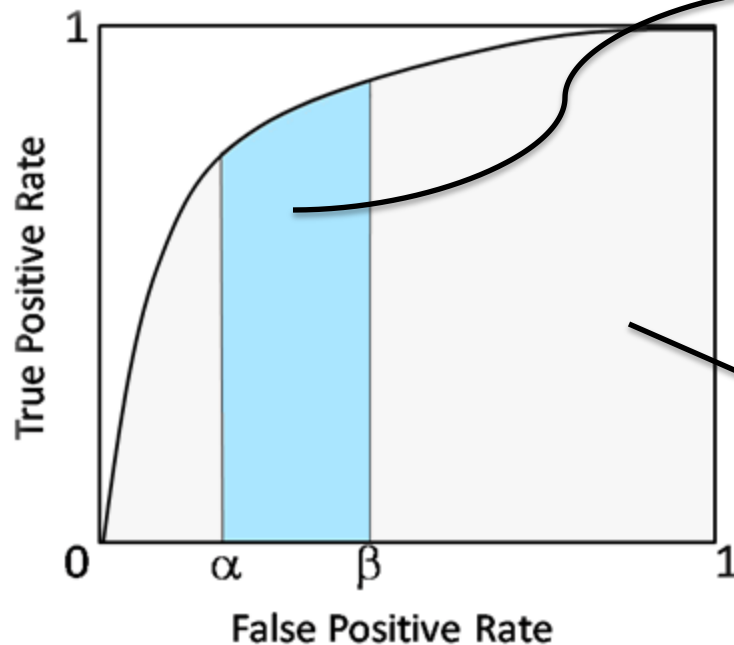
$$\min_{w, \xi \geq 0} \frac{1}{2} \|w\|^2 + C\xi$$

s.t. $\forall \pi \in \Pi_{m,n} :$

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

SVMpAUC (ICML 2013)

Narasimhan, H. and Agarwal, S. “A *structural SVM* based approach for optimizing partial AUC”, ICML 2013.



SVMpAUC

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

s.t. $\forall \pi \in \Pi_{m,n} :$

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

SVM-AUC

Joachims, 2005

Improved Version of SVMpAUC

Tighter upper bound

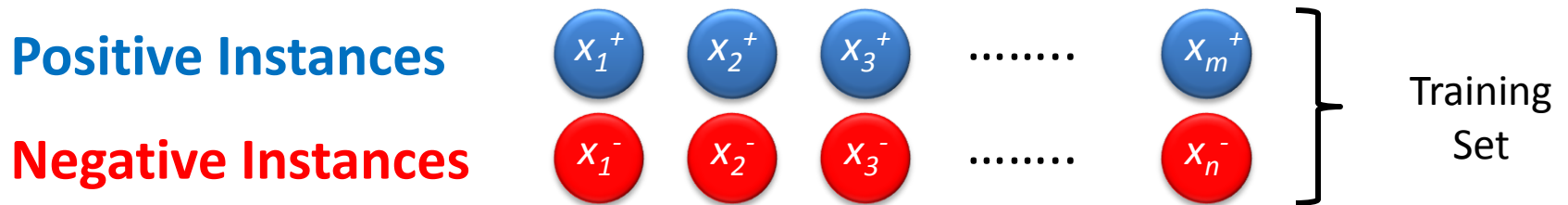
Improved accuracy

Better runtime guarantee

Outline

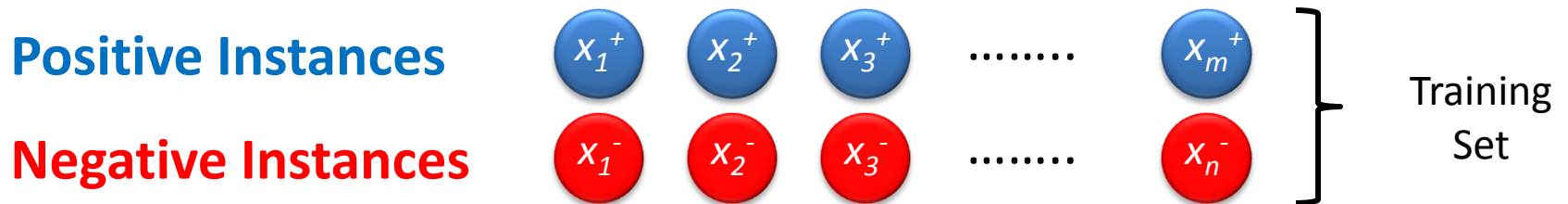
- Overview of SVMpAUC
- Upper Bound Optimized by SVMpAUC
- Improved Formulation: SVMpAUC-tight
- Optimization Methods
- Experiments

Receiver Operating Characteristic Curve



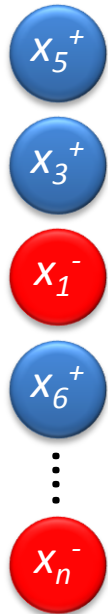
GOAL? Learn a scoring function $f : X \rightarrow \mathbb{R}$

Receiver Operating Characteristic Curve



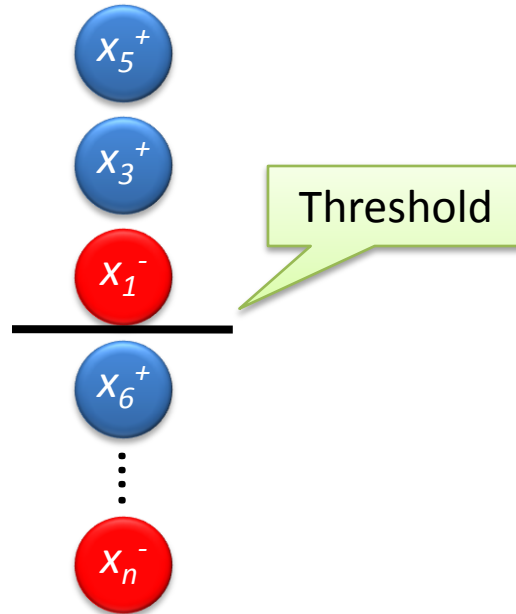
GOAL? Learn a scoring function $f : X \rightarrow \mathbb{R}$

Rank objects

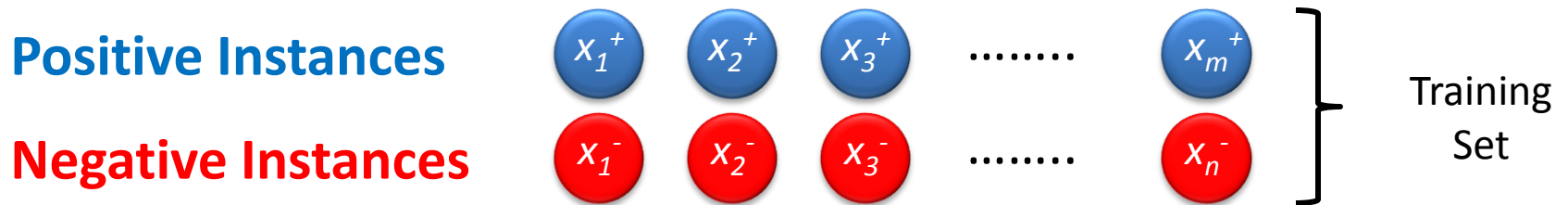


or

Build a classifier

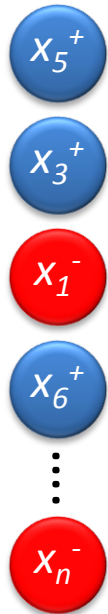


Receiver Operating Characteristic Curve



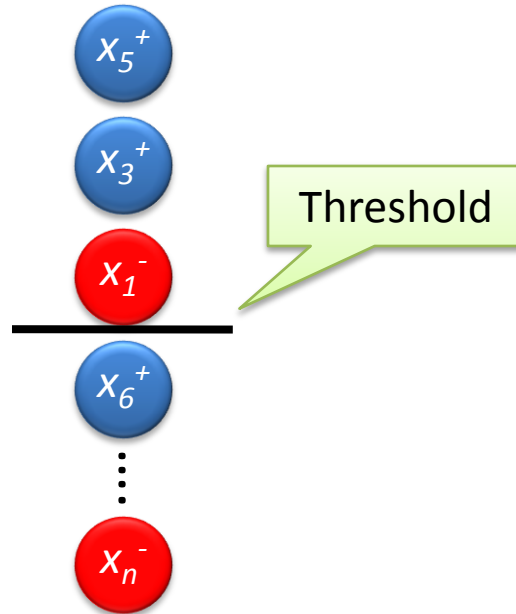
GOAL? Learn a scoring function $f : X \rightarrow \mathbb{R}$

Rank objects

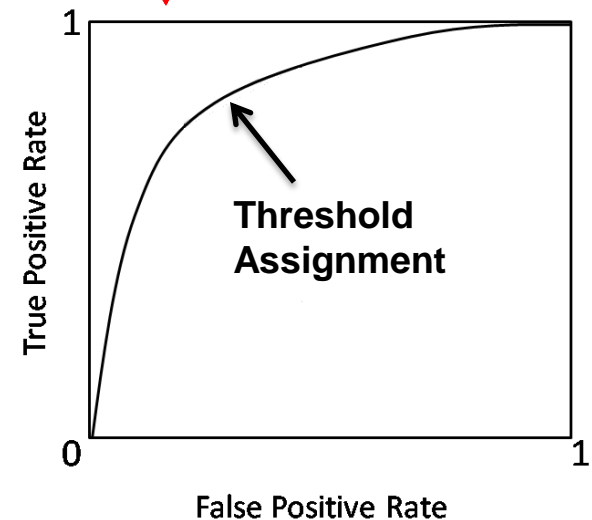


or

Build a classifier

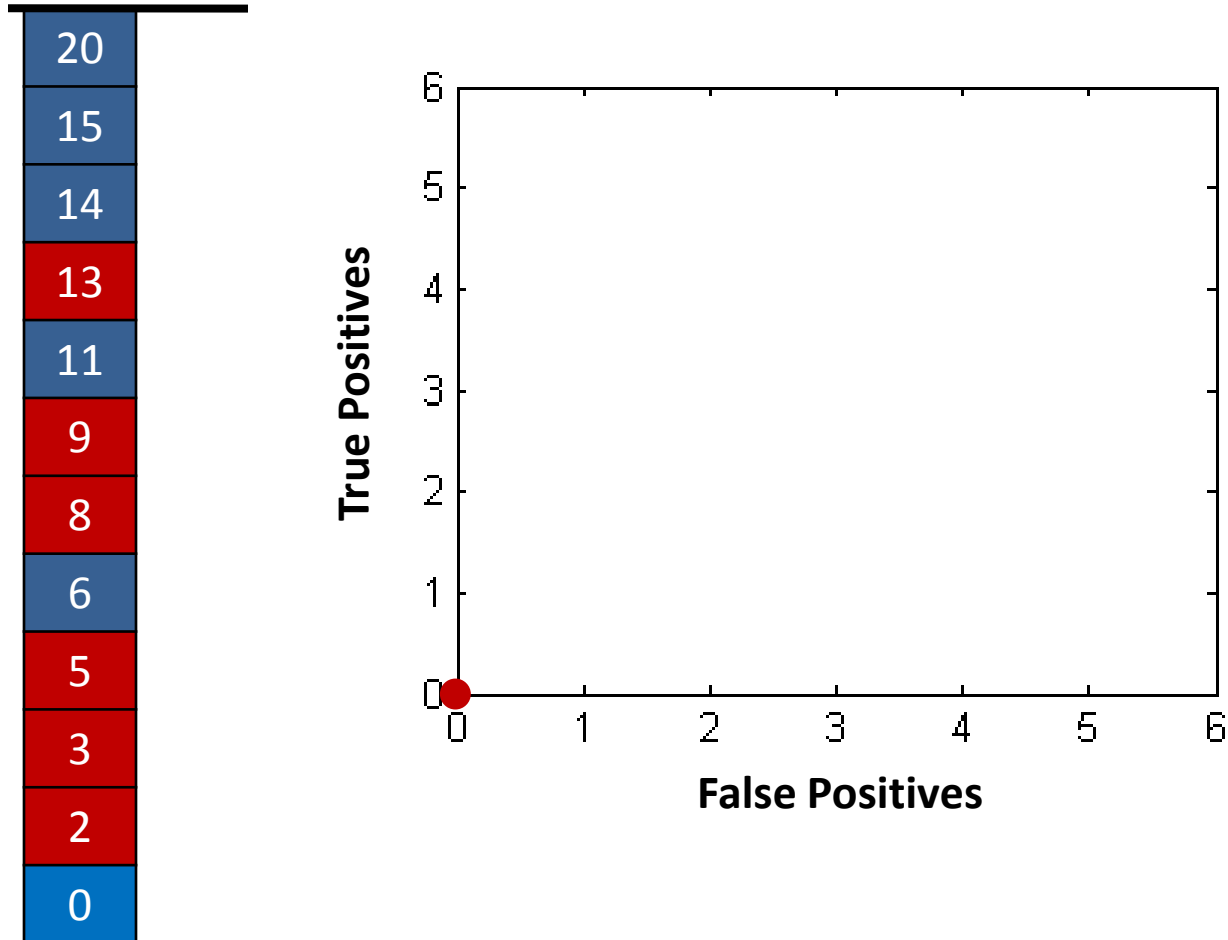


Quality of scoring function?



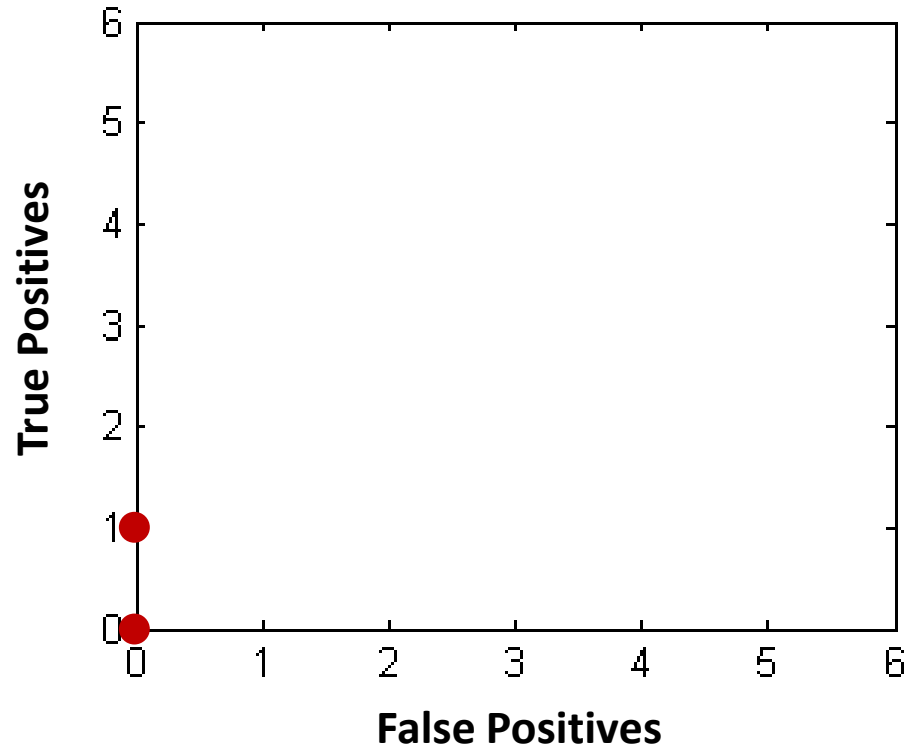
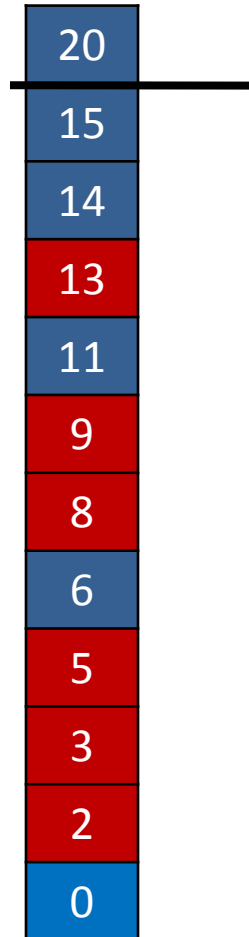
Receiver Operating Characteristic Curve

ROC Curve



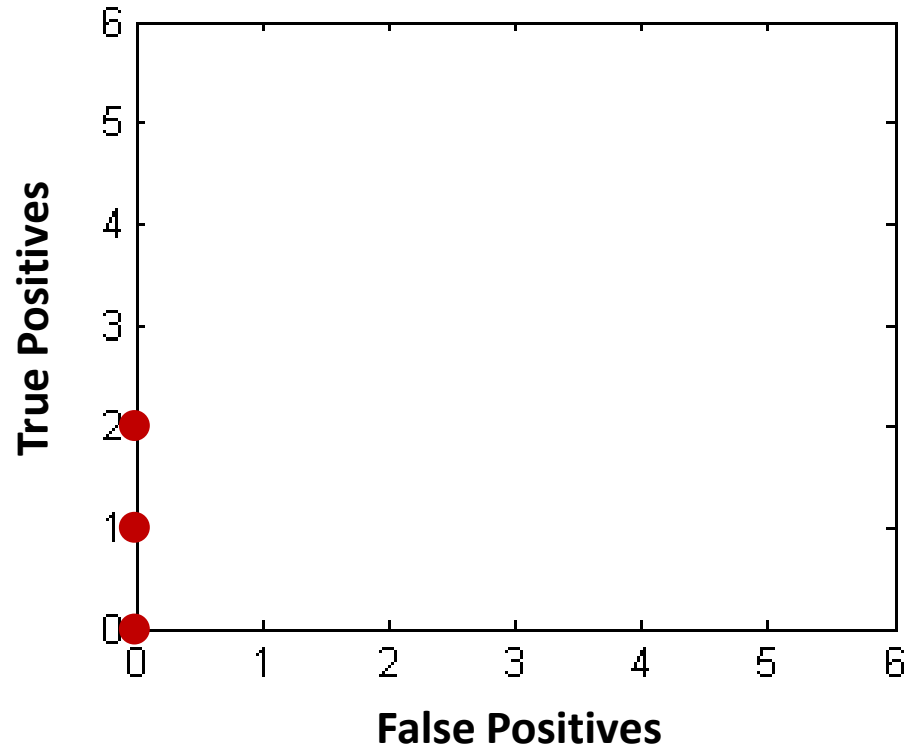
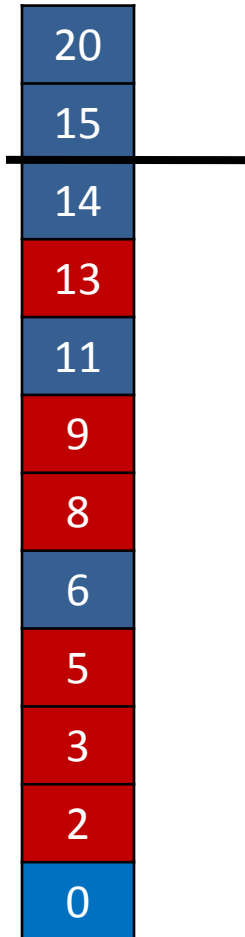
Receiver Operating Characteristic Curve

ROC Curve



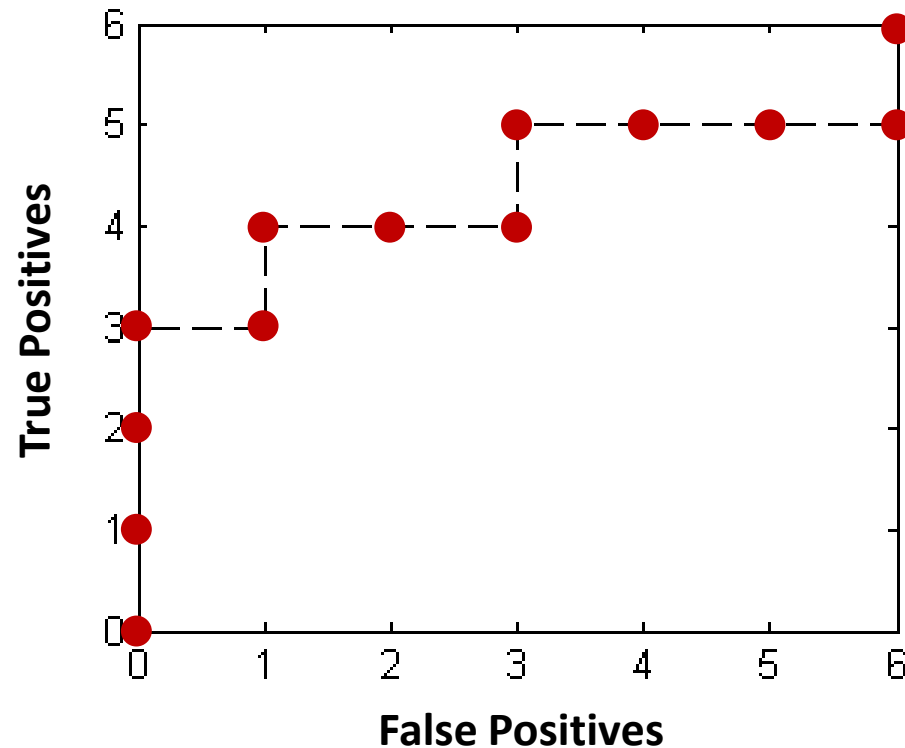
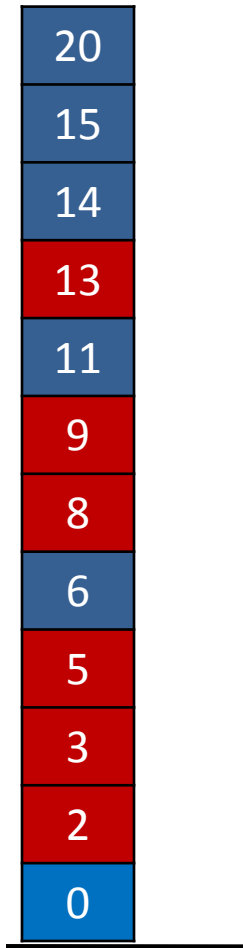
Receiver Operating Characteristic Curve

ROC Curve



Receiver Operating Characteristic Curve

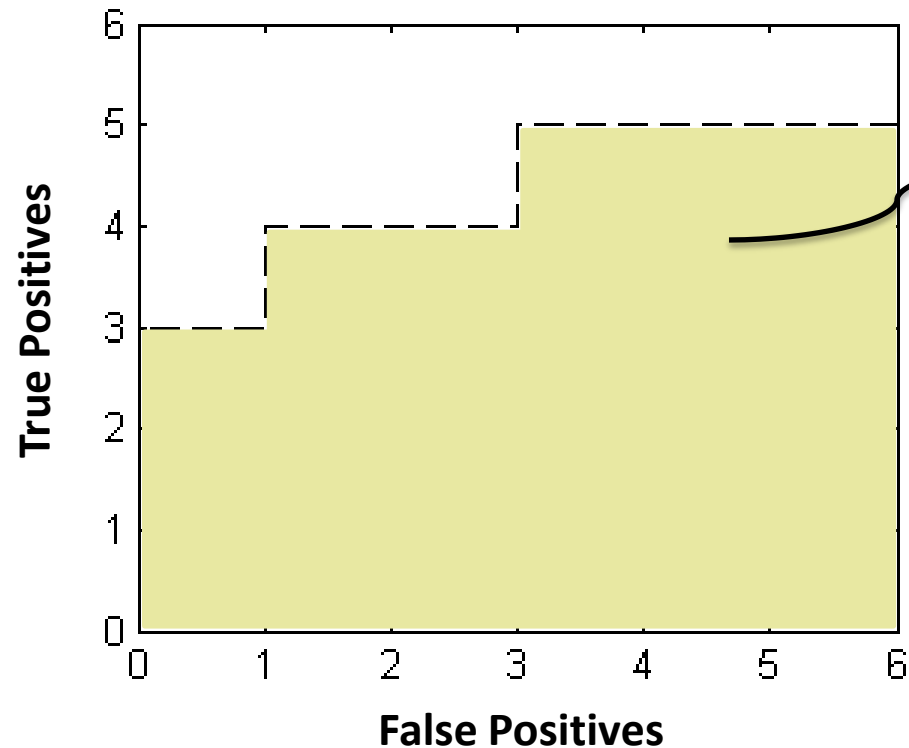
ROC Curve



Receiver Operating Characteristic Curve

ROC Curve

20
15
14
13
11
9
8
6
5
3
2
0

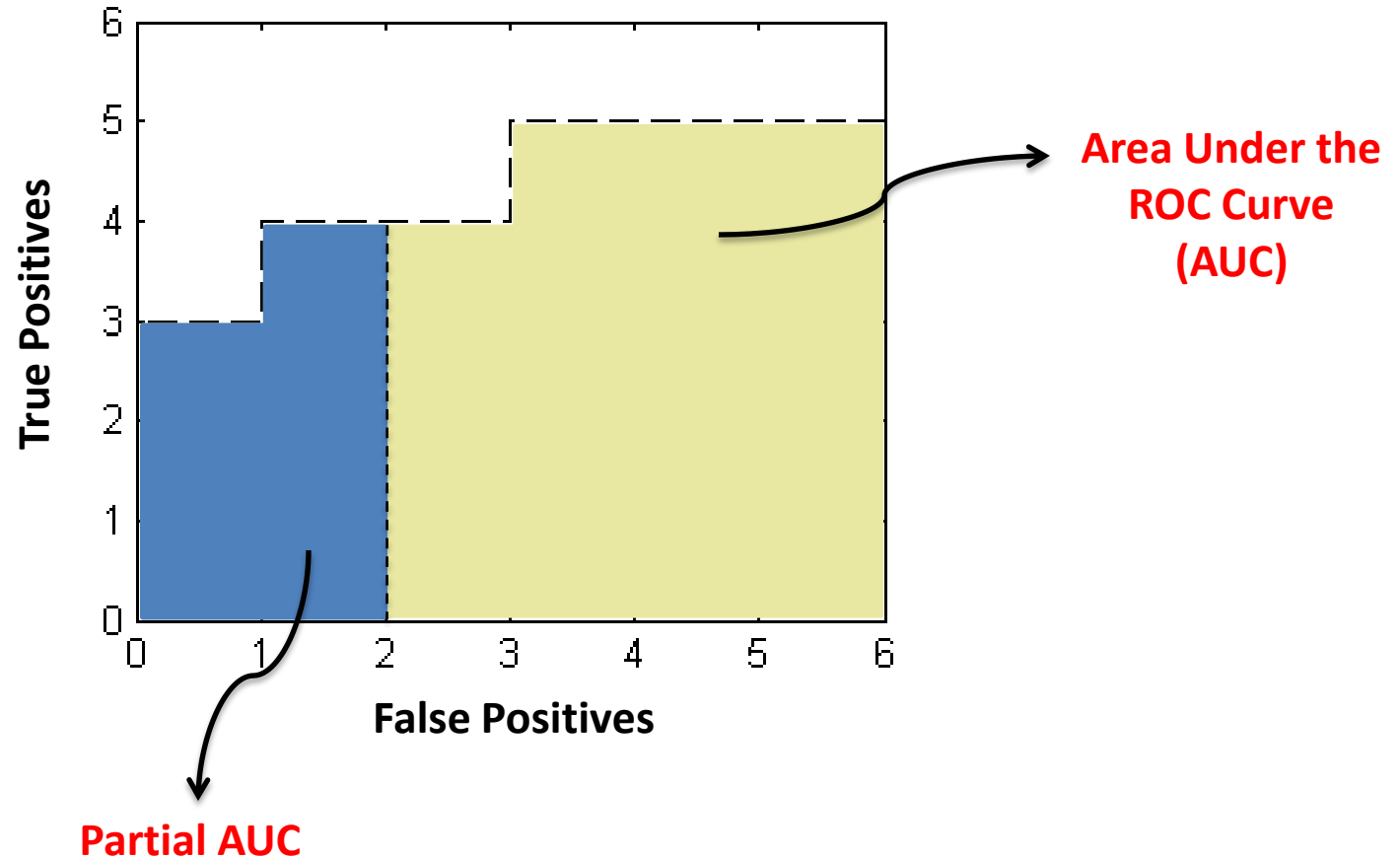


Area Under the
ROC Curve
(AUC)

Receiver Operating Characteristic Curve

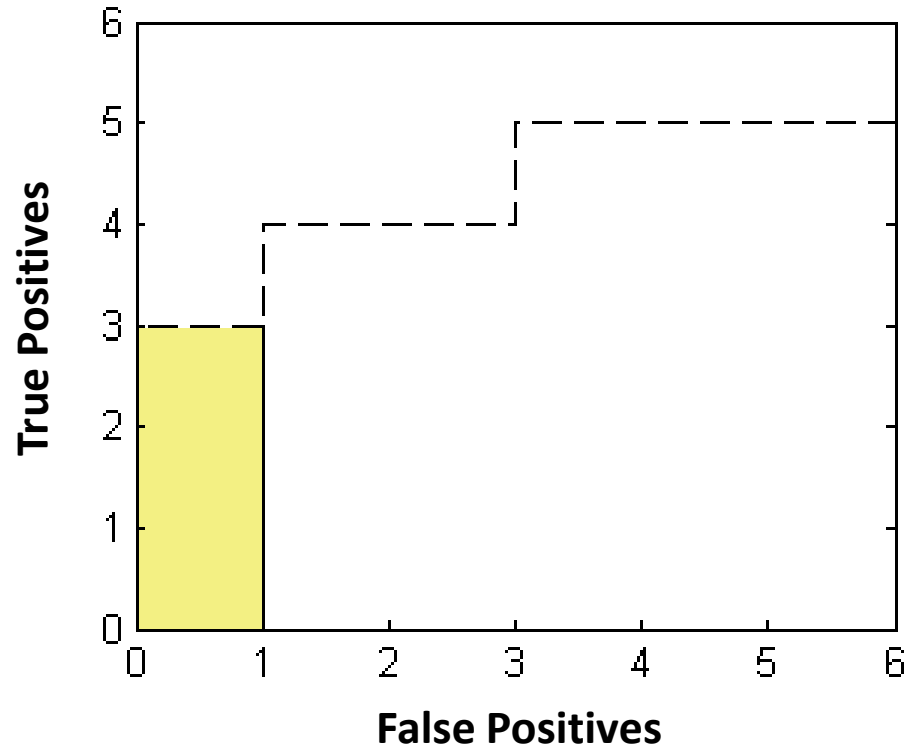
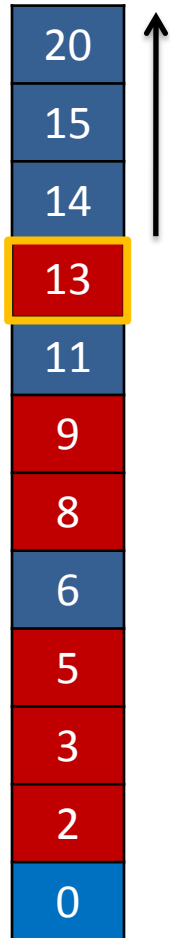
ROC Curve

20
15
14
13
11
9
8
6
5
3
2
0



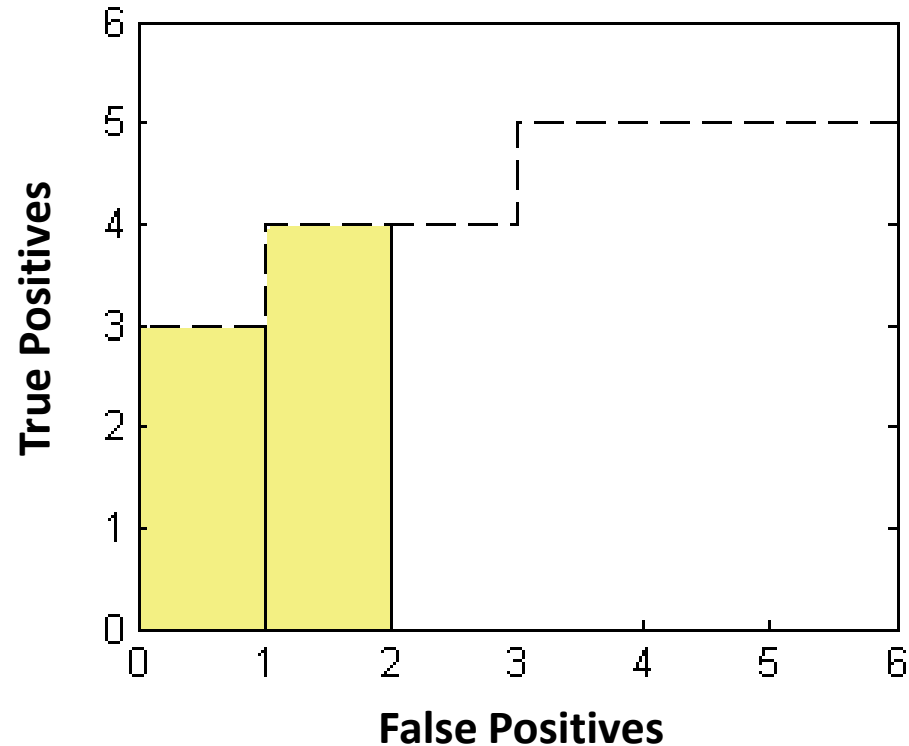
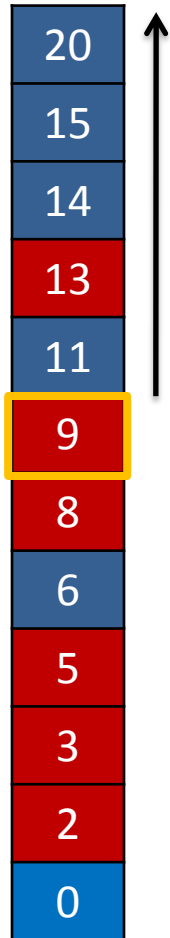
Receiver Operating Characteristic Curve

ROC Curve



Receiver Operating Characteristic Curve

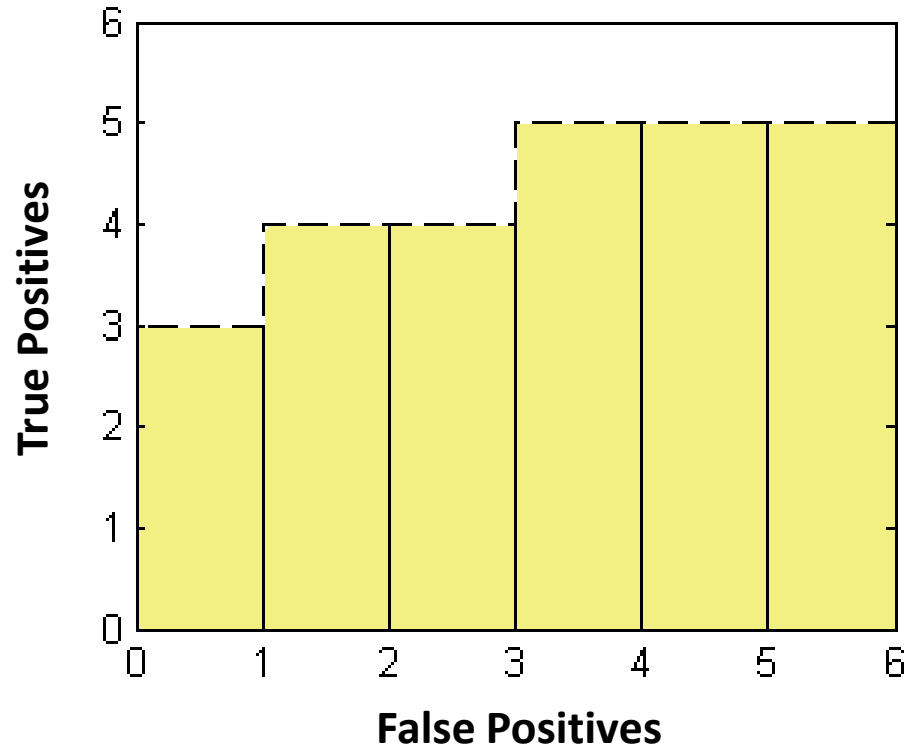
ROC Curve



Receiver Operating Characteristic Curve

ROC Curve

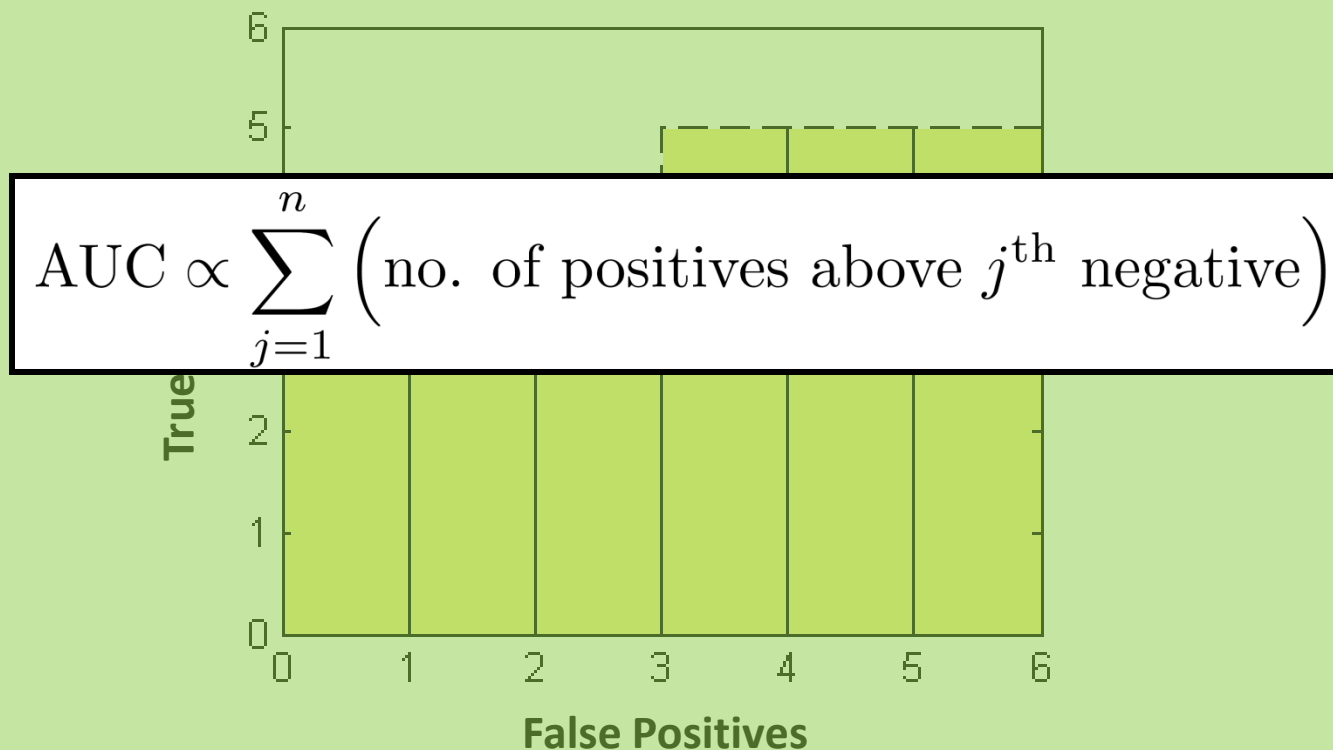
20
15
14
13
11
9
8
6
5
3
2
0



Receiver Operating Characteristic Curve

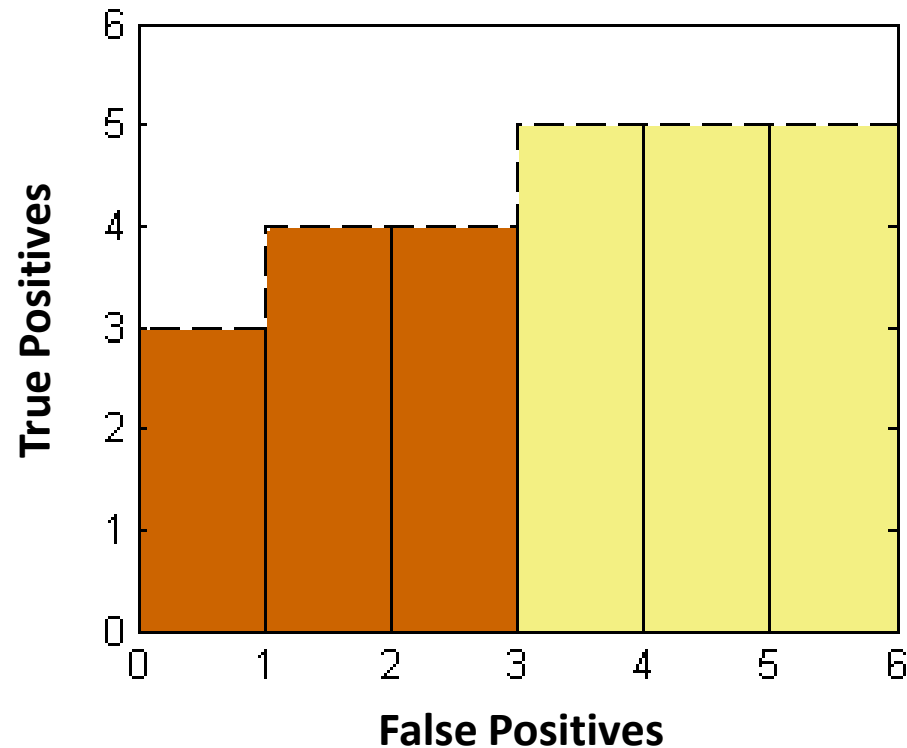
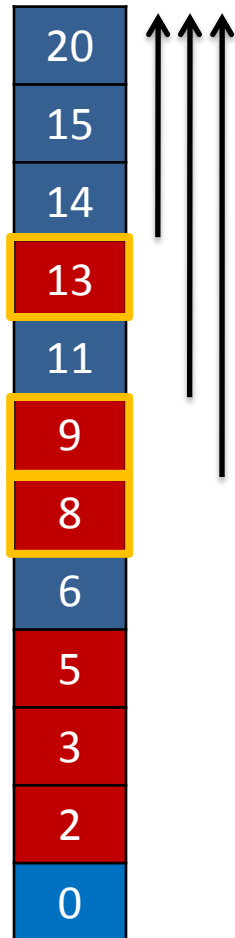
ROC Curve

20
15
14
13
11
9
8
6
5
3
2
0



Receiver Operating Characteristic Curve

ROC Curve

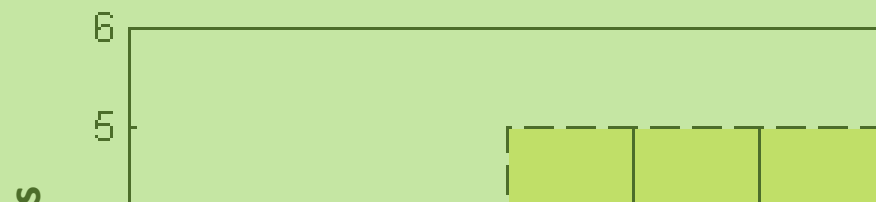
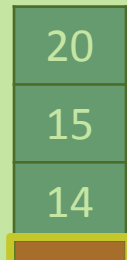


$$\beta = 0.5$$

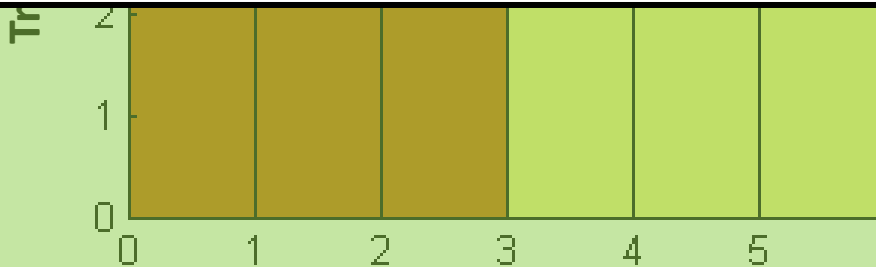
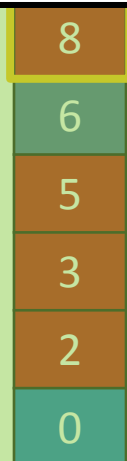
Top 3 negatives!

Receiver Operating Characteristic Curve

ROC Curve



$$\text{pAUC} \propto \sum_{\text{negatives } j \text{ in FPR range } [\alpha, \beta]}^n \left(\text{no. of positives above } j^{\text{th}} \text{ negative} \right)$$

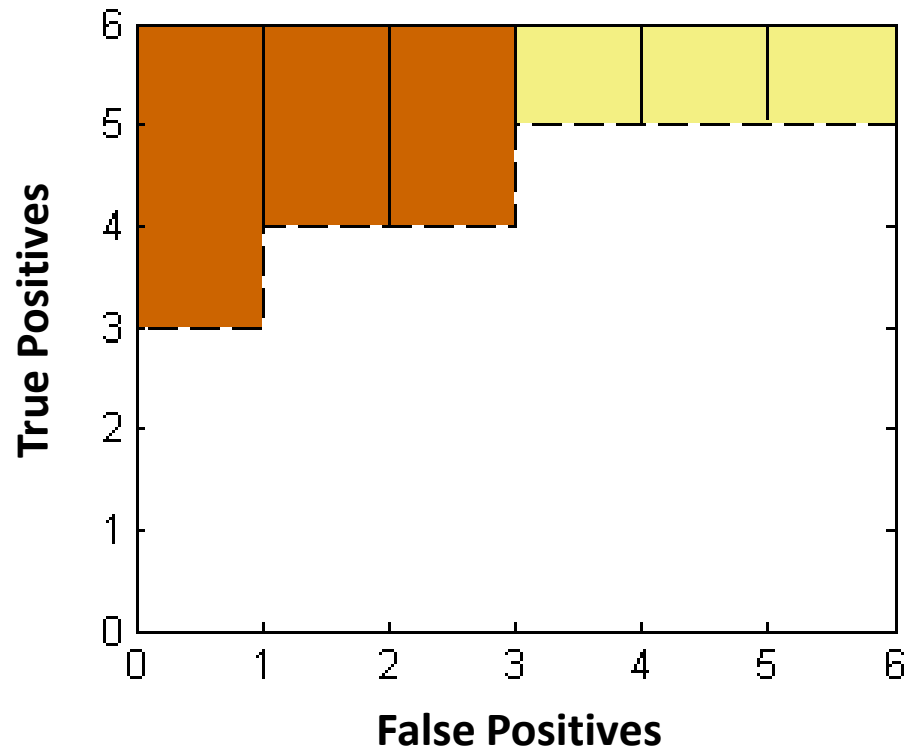


False Positives

Receiver Operating Characteristic Curve

ROC Curve

20
15
14
13
11
9
8
6
5
3
2
0

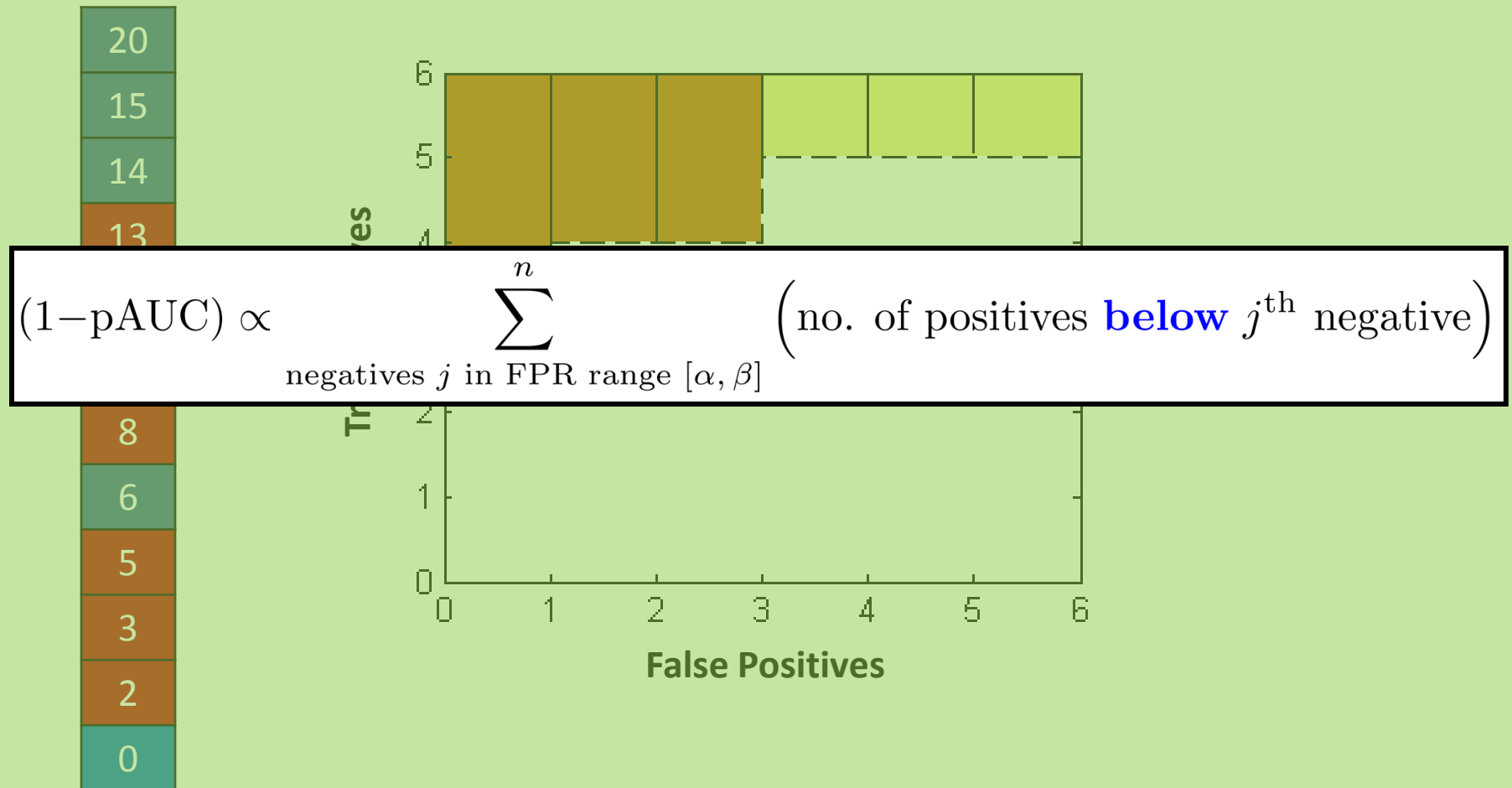


$\beta = 0.5$

Top 3 negatives!

Receiver Operating Characteristic Curve

ROC Curve



$(1 - \text{pAUC})$ for f

Convex Upper Bound

$(1 - \text{pAUC})$ for f



Convex Upper Bound

$(1 - \text{pAUC})$ for f + Regularizer

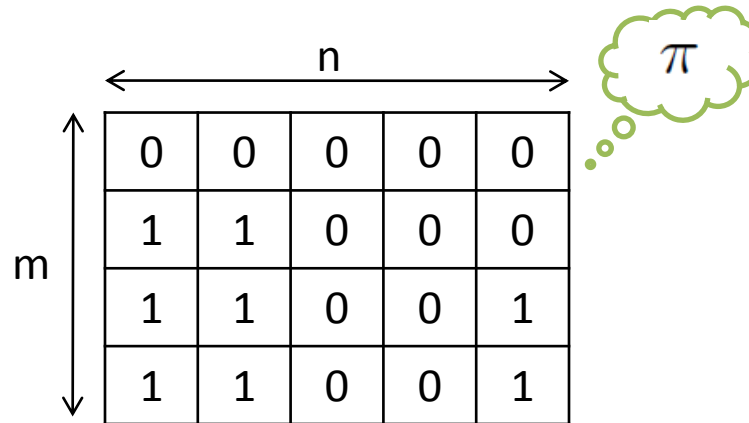


SVM_pAUC (ICML 2013)

SVM_pAUC: Structural SVM Approach
Narasimhan and Agarwal, 2013

SVMpAUC (ICML 2013)

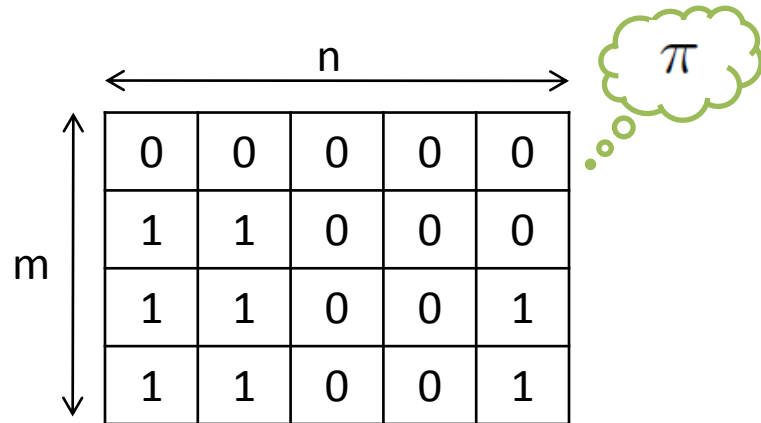
Ordering of
training examples:



SVMpAUC: Structural SVM Approach
Narasimhan and Agarwal, 2013

SVMpAUC (ICML 2013)

Ordering of
training examples:

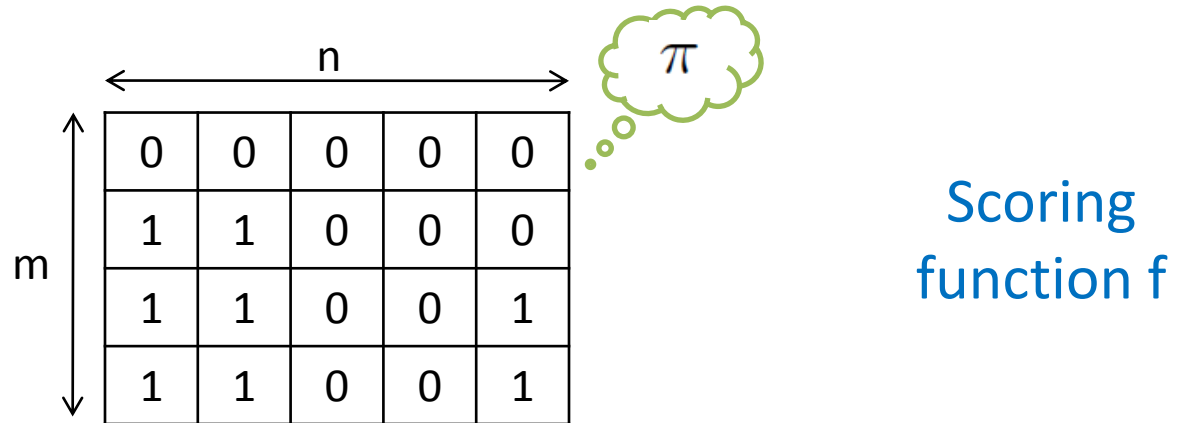


Scoring
function f

SVMpAUC: Structural SVM Approach
Narasimhan and Agarwal, 2013

SVMpAUC (ICML 2013)

Ordering of
training examples:

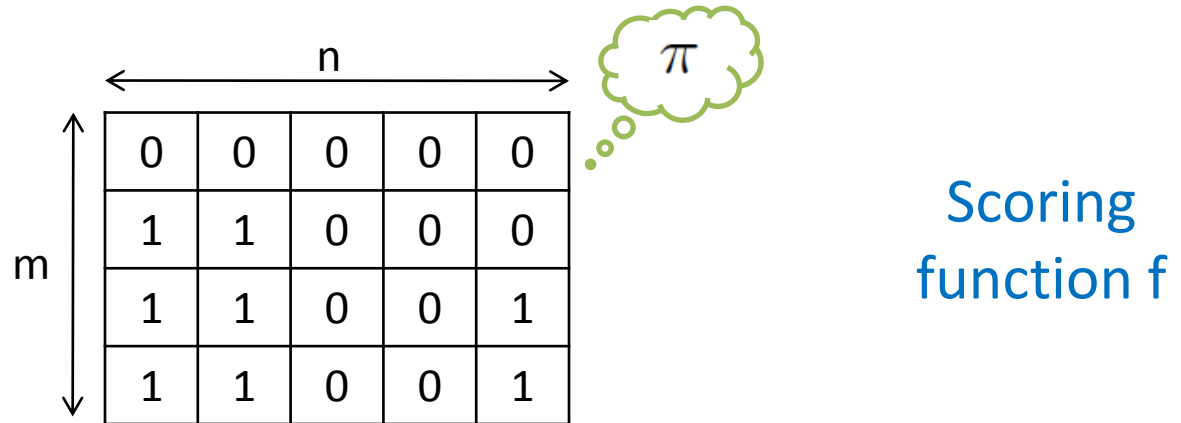


$$(1 - \text{pAUC})_{\text{for } \pi} + \text{term capturing agreement between } \pi \text{ and } f \text{ on all pairs}$$

SVMpAUC: Structural SVM Approach
Narasimhan and Agarwal, 2013

SVMpAUC (ICML 2013)

Ordering of
training examples:



$$\max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times n}} \left(\underset{\text{for } \pi}{(1 - \text{pAUC})} + \underset{\text{on all pairs}}{\text{term capturing agreement between } \pi \text{ and } f} \right)$$

SVMpAUC: Structural SVM Approach
Narasimhan and Agarwal, 2013

Convex Upper Bound

$(1 - \text{pAUC})$ for f + Regularizer

\leq

$$\max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times n}} \left(\begin{array}{l} (1 - \text{pAUC}) \\ \text{for } \pi \end{array} + \begin{array}{l} \text{term capturing} \\ \text{agreement between } \pi \text{ and } f \\ \text{on all pairs} \end{array} \right)$$

Convex Upper Bound

$(1 - \text{pAUC})$ for f + Regularizer

\leq

$$\max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times n}} \left(\underset{\text{for } \pi}{(1 - \text{pAUC})} + \underset{\substack{\text{term capturing} \\ \text{agreement between } \pi \text{ and } f \\ \text{on all pairs}}}{\phantom{(1 - \text{pAUC})}} \right)$$

How does this upper bound look?

Convex Upper Bound

$(1 - \text{pAUC})$ for f + Regularizer

\leq

$$\max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times n}} \left(\begin{array}{l} (1 - \text{pAUC}) \\ \text{for } \pi \end{array} + \begin{array}{l} \text{term capturing} \\ \text{agreement between } \pi \text{ and } f \\ \text{on all pairs} \end{array} \right)$$

Can we obtain a tighter upper bound?

Outline

- Overview of SVMpAUC
- Upper Bound Optimized by SVMpAUC
- Improved Formulation: SVMpAUC-tight
- Optimization Methods
- Experiments

Upper bound we want?

1 - pAUC

$$\propto \sum_{\substack{\text{negatives } j \text{ in} \\ \text{FPR range } [\alpha, \beta]}}^n \left(\text{no. of positives } \textcolor{blue}{\text{below}} j^{\text{th}} \text{ negative} \right)$$

Upper bound we want?

1 - pAUC

$$\propto \sum_{\substack{\text{negatives } j \text{ in} \\ \text{FPR range } [\alpha, \beta]}}^n \sum_{i=1}^m \mathbf{1}(i^{\text{th}} \text{ positive ranked by } f \text{ below } j^{\text{th}} \text{ negative?})$$

Upper bound we want?

1 - pAUC

$$\propto \sum_{\substack{\text{negatives } j \text{ in} \\ \text{FPR range } [\alpha, \beta]}}^n \sum_{i=1}^m \mathbf{1}(f(x_i^+) - f(x_j^-) \leq 0)$$

Upper bound we want?

1 - pAUC

$$\propto \sum_{\substack{\text{negatives } j \text{ in} \\ \text{FPR range } [\alpha, \beta]}}^n \sum_{i=1}^m \mathbf{1}(f(x_i^+) - f(x_j^-) \leq 0)$$

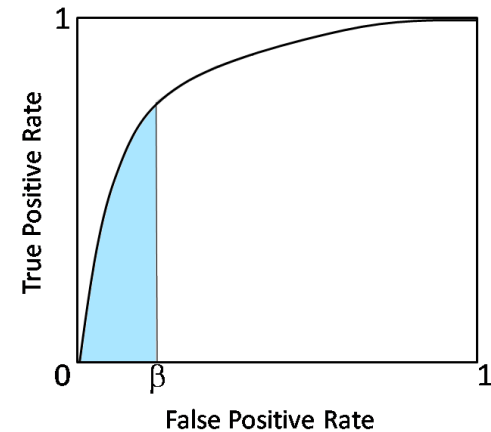
$$\leq \sum_{\substack{\text{negatives } j \text{ in} \\ \text{FPR range } [\alpha, \beta]}}^n \sum_{i=1}^m \text{hinge-loss}(f(x_i^+) - f(x_j^-))$$

pair-wise hinge loss!

Upper optimized by SVMpAUC?

$$\max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times n}} \left(\underset{\text{for } \pi}{(1 - \text{pAUC})} + \underset{\text{on all pairs}}{\overset{\text{term capturing}}{\text{agreement between } \pi \text{ and } f}} \right)$$

Upper optimized by SVMpAUC?



$$\begin{aligned}
 & \max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times n}} \left(\underbrace{(1 - \text{pAUC})}_{\text{for } \pi} + \underbrace{\text{term capturing agreement between } \pi \text{ and } f}_{\text{on all pairs}} \right) \\
 & \qquad \qquad \qquad = \\
 & \text{pair-wise hinge loss} \quad + \quad \text{extra term}
 \end{aligned}$$

Upper optimized by SVMpAUC?



Subset of pairs of
positive-negative
examples

$$\max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times n}} \left(\underset{\text{for } \pi}{(1 - \text{pAUC})} + \underset{\substack{\text{term capturing} \\ \text{agreement between } \pi \text{ and } f \\ \text{on } \underline{\text{all pairs}}}}{\quad} \right)$$

=

pair-wise hinge loss + extra term

Upper optimized by SVMpAUC?



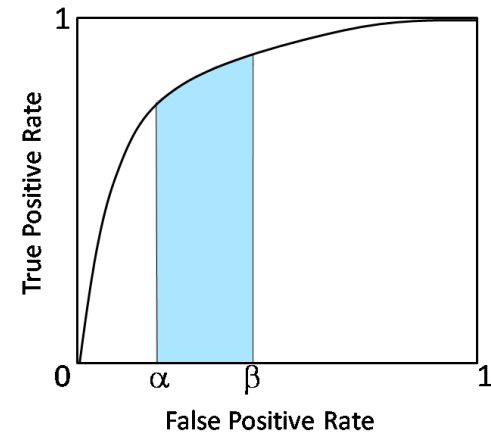
Subset of pairs of
positive-negative
examples

$$\max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times n}} \left(\underset{\text{for } \pi}{(1 - \text{pAUC})} + \underset{\substack{\text{term capturing} \\ \text{agreement between } \pi \text{ and } f \\ \text{on } \underline{\text{all pairs}}}}{\quad} \right)$$

=

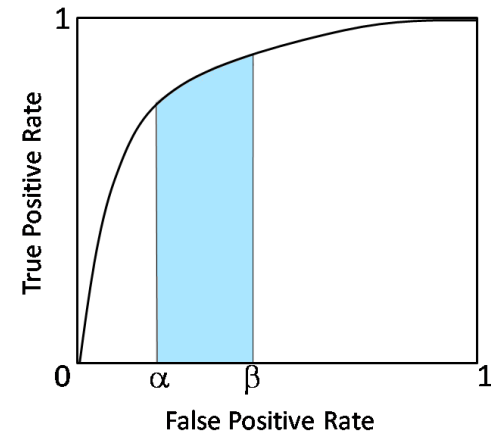
pair-wise hinge loss + ~~extra term~~ ?

Upper optimized by SVMpAUC?



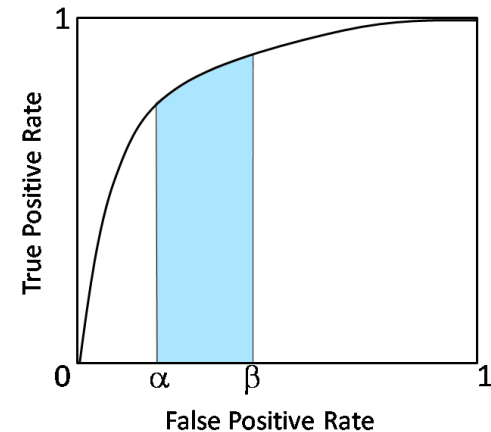
$$\max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times n}} \left(\begin{array}{c} (1 - \text{pAUC}) \\ \text{for } \pi \end{array} + \begin{array}{c} \text{term capturing} \\ \text{agreement between } \pi \text{ and } f \\ \text{on all pairs} \end{array} \right)$$

Upper optimized by SVMpAUC?



$$\begin{aligned}
 & \max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times n}} \left(\underbrace{(1 - \text{pAUC})}_{\text{for } \pi} + \underbrace{\text{term capturing agreement between } \pi \text{ and } f}_{\text{on all pairs}} \right) \\
 & \leq \\
 & \text{pair-wise hinge loss} + \text{extra term}
 \end{aligned}$$

Upper optimized by SVMpAUC?



approx. pair-wise hinge loss + extra term

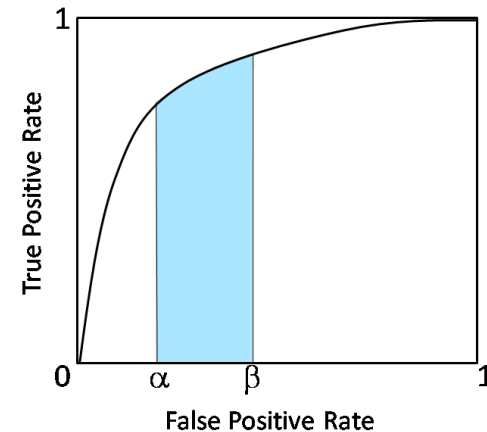
\leq

$$\max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times n}} \left(\underset{\text{for } \pi}{(1 - \text{pAUC})} + \underset{\substack{\text{term capturing} \\ \text{agreement between } \pi \text{ and } f \\ \text{on all pairs}}}{\phantom{(1 - \text{pAUC})}} \right)$$

\leq

pair-wise hinge loss + extra term

Upper optimized by SVMpAUC?



approx. pair-wise hinge loss + ~~extra term~~?

\leq

$$\max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times n}} \left(\underset{\text{for } \pi}{(1 - \text{pAUC})} + \underset{\substack{\text{term capturing} \\ \text{agreement between } \pi \text{ and } f \\ \text{on all pairs}}}{\phantom{(1 - \text{pAUC})}} \right)$$

\leq

pair-wise hinge loss + ~~extra term~~?

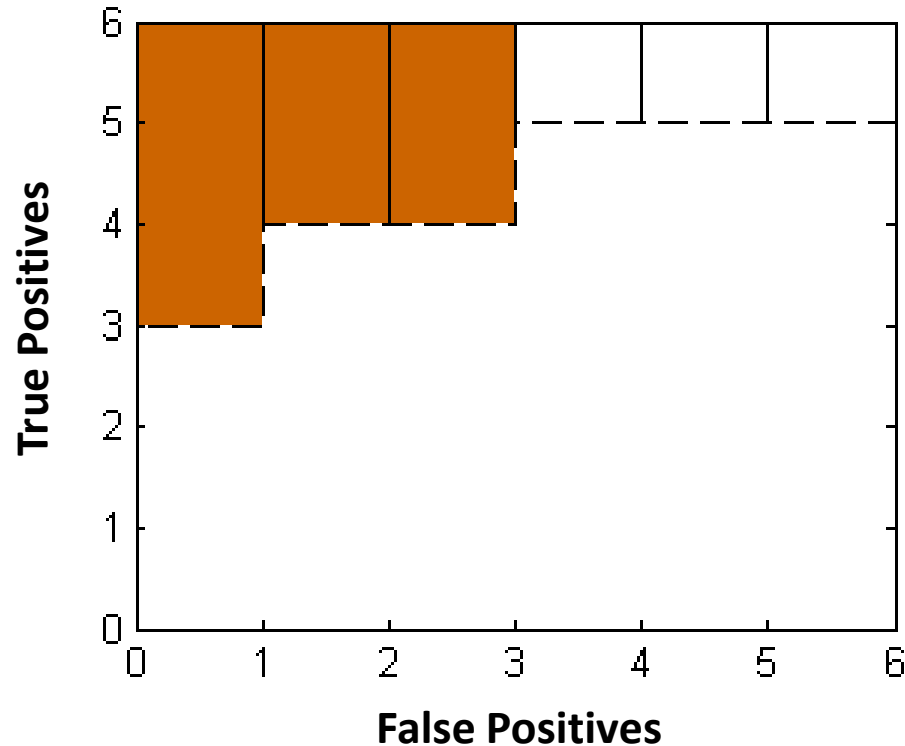
Outline

- Overview of SVMpAUC
- Upper Bound Optimized by SVMpAUC
- Improved Formulation: SVMpAUC-tight
- Optimization Methods
- Experiments

Rewriting the Partial AUC Loss

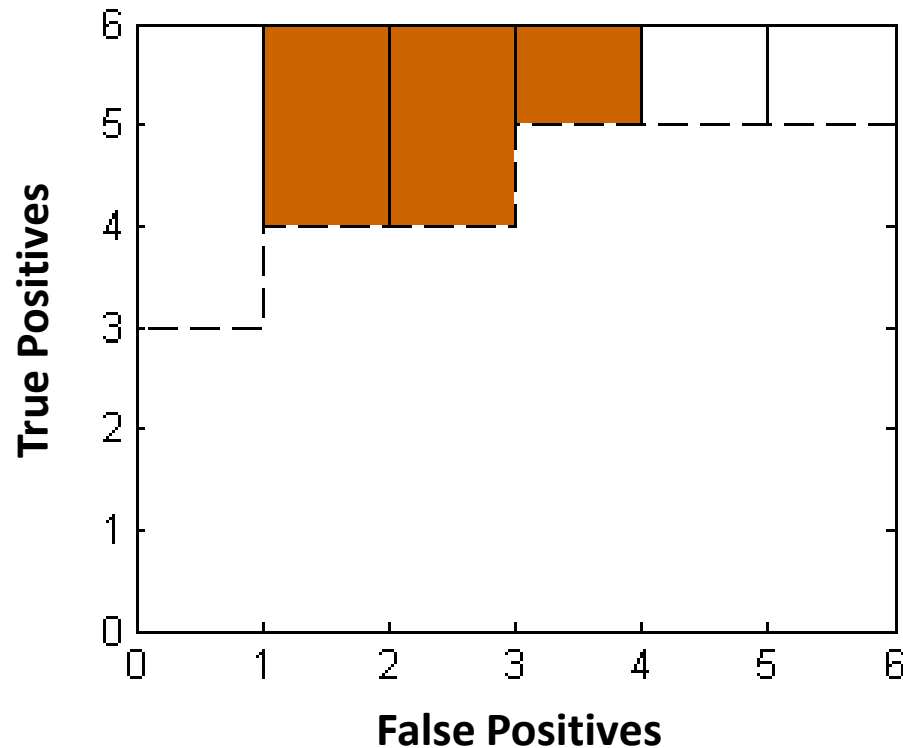
$$\alpha = 0, \quad \beta = 0.5$$

$$3 + 2 + 2 = 7$$



Rewriting the Partial AUC Loss

$$\alpha = 0, \quad \beta = 0.5$$

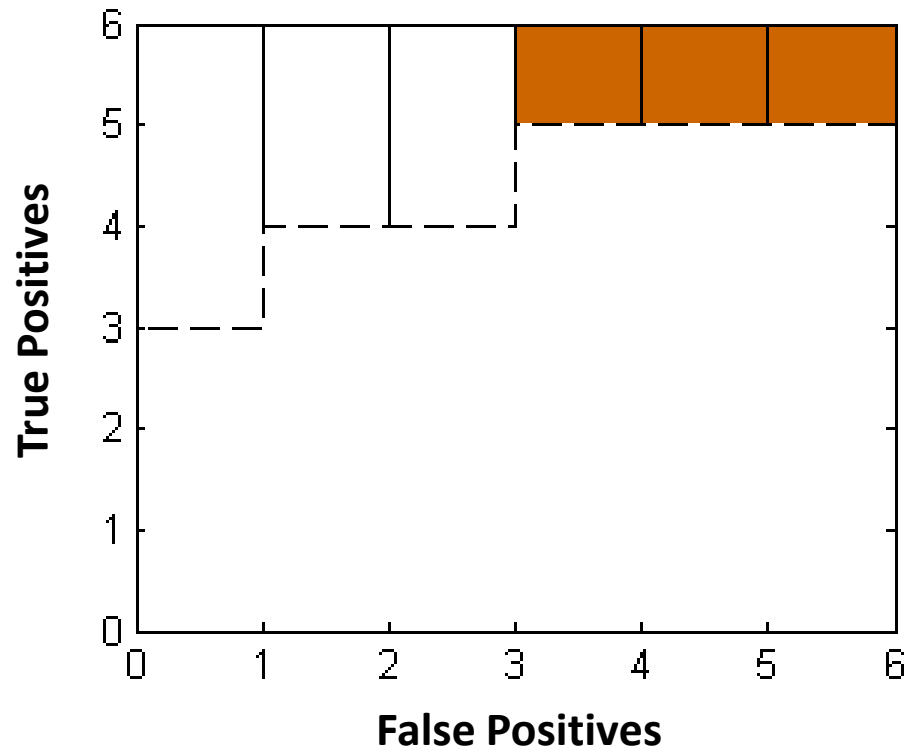
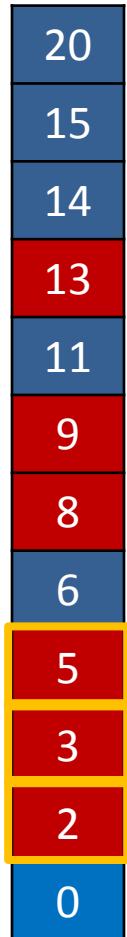


$$3 + 2 + 2 = 7$$

$$2 + 2 + 1 = 5$$

Rewriting the Partial AUC Loss

$$\alpha = 0, \quad \beta = 0.5$$



$$3 + 2 + 2 = 7$$

$$2 + 2 + 1 = 5$$

.

.

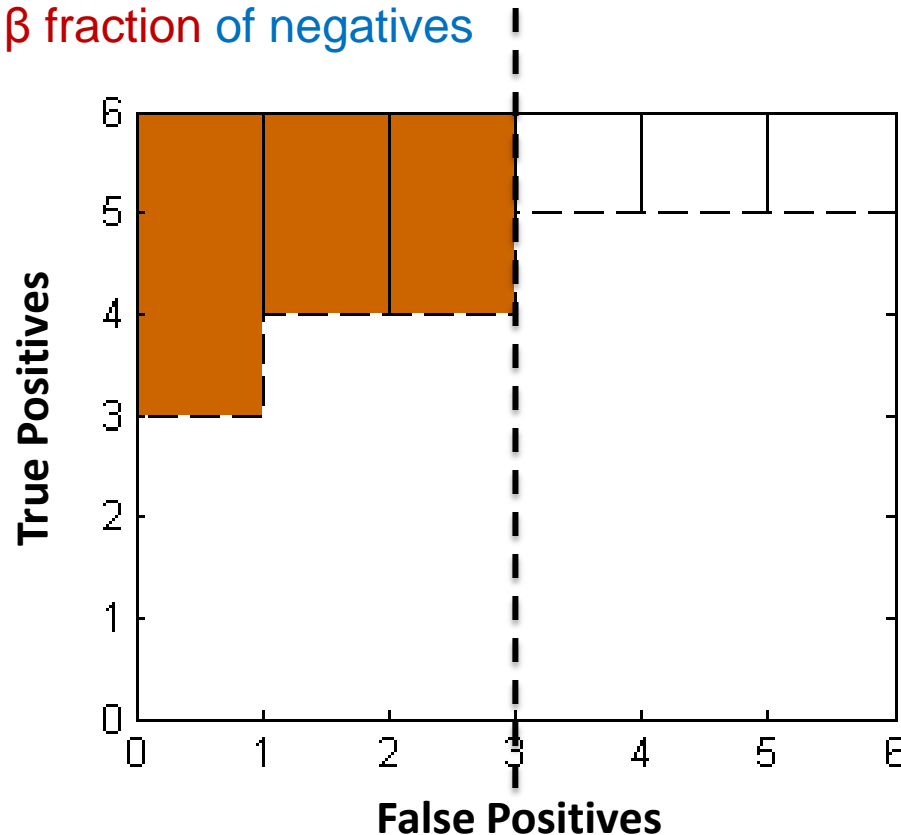
.

$$1 + 1 + 1 = 3$$

Rewriting the Partial AUC Loss

1 - AUC restricted to
top β fraction of negatives

$$\alpha = 0, \quad \beta = 0.5$$



Maximum!

$$3 + 2 + 2 = 7$$

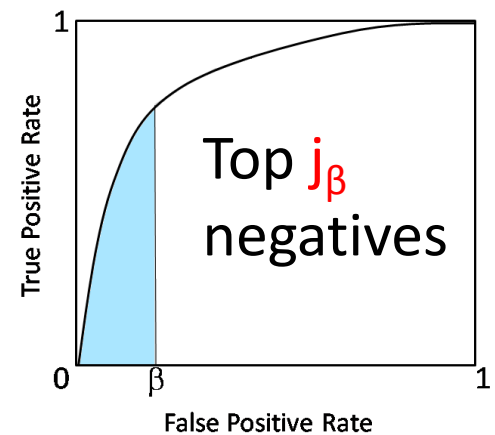
$$2 + 2 + 1 = 5$$

.

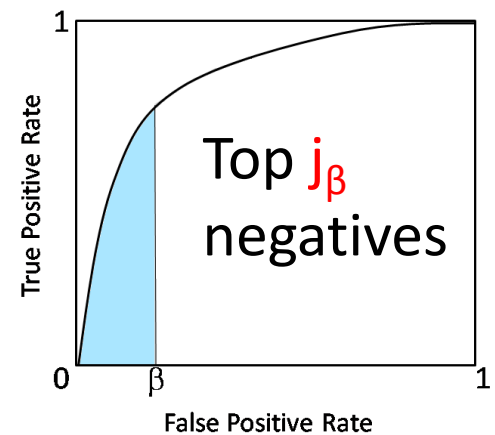
.

.

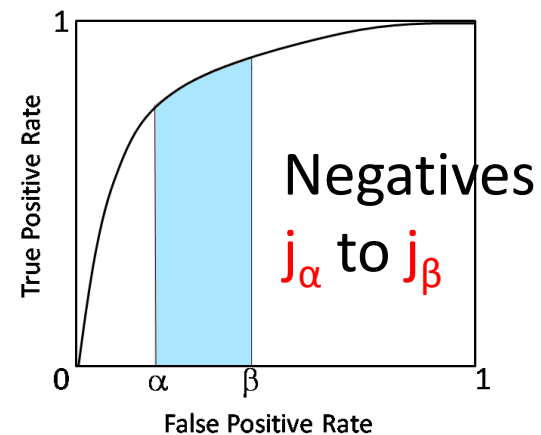
$$1 + 1 + 1 = 3$$



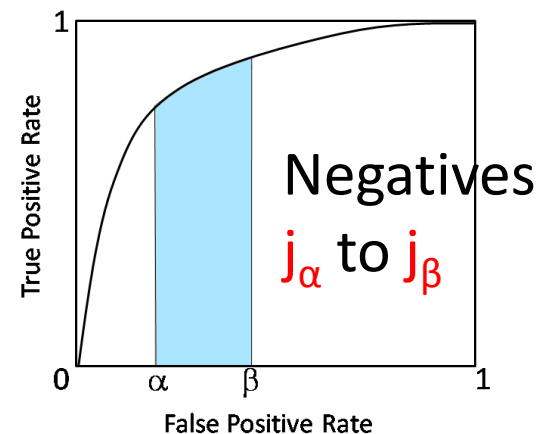
$$(1 - \text{pAUC}) \propto \max_{\substack{\text{subsets } S \text{ of negatives} \\ \text{of size } j_\beta}} \left(1 - \text{AUC restricted to negatives in } S \right)$$



$$(1 - \text{pAUC}) \propto \max_{\substack{\text{subsets } S \text{ of negatives} \\ \text{of size } j_\beta}} \left(\frac{\text{SVM-AUC}}{1 - \text{AUC restricted to negatives in } S} \right)$$



$$(1-\text{pAUC}) \propto \max_{\substack{\text{subsets } S \text{ of negatives} \\ \text{of size } j_\beta}} \left(1-\text{pAUC} \text{ restricted to negatives in } S \right)$$



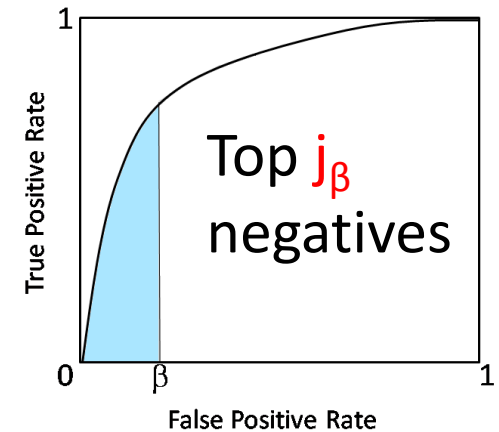
$$(1-\text{pAUC}) \propto \max_{\substack{\text{subsets } S \text{ of negatives} \\ \text{of size } j_\beta}} \left(\text{Truncated SVMpAUC} \right)$$

$1-\text{pAUC}$ restricted to negatives in S

SVMpAUC-tight: Improved Formulation

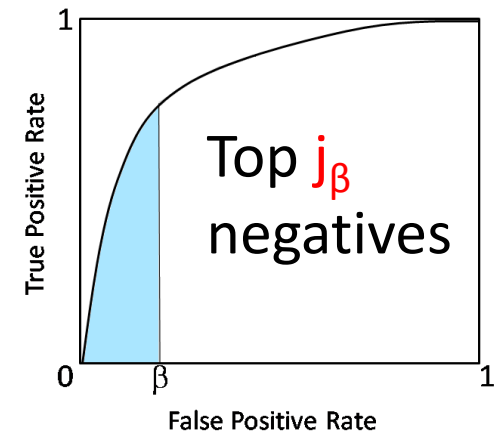
$$\max_{\substack{\text{subsets } S \text{ of negatives} \\ \text{of size } j_\beta}} \left[\text{SVMpAUC objective restricted to } S \right]$$

SVMpAUC-tight: Improved Formulation



$$\max_{\substack{\text{subsets } S \text{ of negatives} \\ \text{of size } j_\beta}} \left[\max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times j_\beta}} \left((1 - \text{AUC})_{\text{for } \pi} + \begin{array}{l} \text{term capturing} \\ \text{agreement between } \pi \text{ and } f \\ \text{on pairs corresponding to } S \end{array} \right) \right]$$

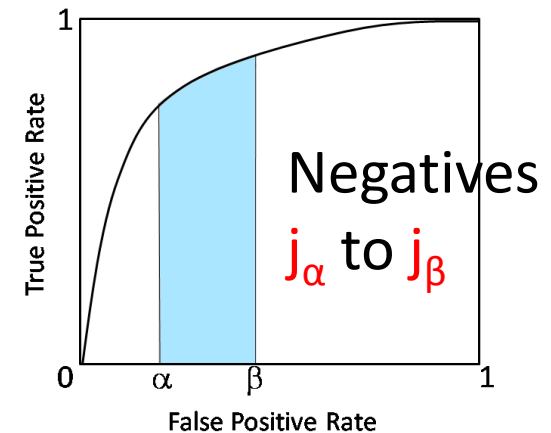
SVMpAUC-tight: Improved Formulation



$$\begin{aligned}
 & \max_{\substack{\text{subsets } S \text{ of negatives} \\ \text{of size } j_\beta}} \left[\max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times j_\beta}} \left((1 - \text{AUC})_{\text{for } \pi} + \text{term capturing agreement between } \pi \text{ and } f \text{ on pairs corresponding to } S \right) \right] \\
 & = \\
 & \text{pair-wise hinge loss} + \text{extra term}
 \end{aligned}$$

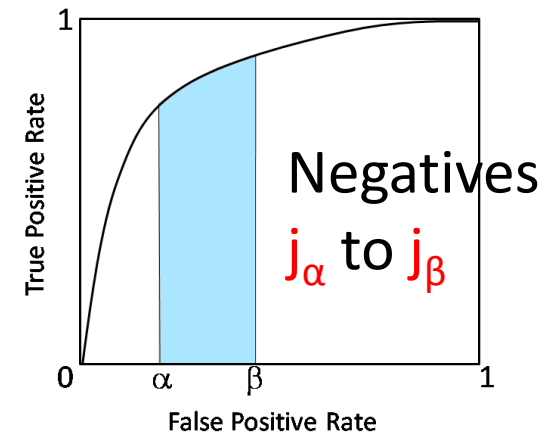
Same pairs of positive-negative examples

SVMpAUC-tight: Improved Formulation



$$\max_{\substack{\text{subsets } S \text{ of negatives} \\ \text{of size } j_\beta}} \left[\max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times j_\beta}} \left(\begin{array}{c} \text{restricted} \\ (1 - \text{pAUC}) \\ \text{for } \pi \end{array} + \begin{array}{c} \text{term capturing} \\ \text{agreement between } \pi \text{ and } f \\ \text{on pairs corresponding to } S \end{array} \right) \right]$$

SVMpAUC-tight: Improved Formulation



approx. pair-wise hinge loss + ~~extra term~~

\leq

$$\max_{\substack{\text{subsets } S \text{ of negatives} \\ \text{of size } j_\beta}} \left[\max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times j_\beta}} \left(\begin{array}{c} \text{restricted} \\ (1 - \text{pAUC}) \\ \text{for } \pi \end{array} + \begin{array}{c} \text{term capturing} \\ \text{agreement between } \pi \text{ and } f \\ \text{on pairs corresponding to } S \end{array} \right) \right]$$

\leq

pair-wise hinge loss + ~~extra term~~

Outline


- Overview of SVMpAUC
- Upper Bound Optimized by SVMpAUC
- Improved Formulation: SVMpAUC-tight
- Optimization Methods
- Experiments

SVMpAUC-tight: Optimization Problem

$$\begin{array}{c}
 \begin{array}{cc}
 \max_{\substack{\text{subsets } S \text{ of negatives} \\ \text{of size } j_\beta}} & \max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times j_\beta}}
 \end{array}
 \left(\begin{array}{c} \text{restricted} \\ (1 - \text{pAUC}) \\ \text{for } \pi \end{array} + \begin{array}{c} \text{term capturing} \\ \text{agreement between } \pi \text{ and } f \\ \text{on pairs corresponding to } S \end{array} \right) \\
 \swarrow \quad \searrow \\
 \text{exponential in size}
 \end{array}
 + \text{Regularizer}$$

SVMpAUC-tight: Optimization Problem

$$\max_{\substack{\text{subsets } S \text{ of negatives} \\ \text{of size } j_\beta}} \max_{\substack{\text{ordering matrices } \pi \\ \text{of size } m \times j_\beta}} \left(\begin{array}{c} \text{restricted} \\ (1 - \text{pAUC}) \\ \text{for } \pi \end{array} + \begin{array}{c} \text{term capturing} \\ \text{agreement between } \pi \text{ and } f \\ \text{on pairs corresponding to } S \end{array} \right) + \text{Regularizer}$$


 exponential in size

$$\begin{aligned} & \min_{w, \xi \geq 0} \frac{1}{2} \|w\|_2^2 + C\xi \\ \text{s.t. } & \forall z \in \mathcal{Z}_\beta, \pi \in \Pi_{m, j_\beta} : \\ & w^\top (\phi_z(S, \pi^*) - \phi_z(S, \pi)) \geq \Delta_\beta(\pi^*, \pi) - \xi \end{aligned}$$

Quadratic program with an exponential number of constraints

SVMpAUC-tight: Cutting-Plane Solver

Repeat:

$$\min_{w, \xi \geq 0} \frac{1}{2} \|w\|_2^2 + C\xi$$

s.t. $\forall z \in \mathcal{Z}_\beta, \pi \in \Pi_{m, j_\beta} :$

$$w^\top (\phi_z(S, \pi^*) - \phi_z(S, \pi)) \geq \Delta_\beta(\pi^*, \pi) - \xi$$

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

SVMpAUC-tight: Cutting-Plane Solver

Repeat:

$$\min_{w, \xi \geq 0} \frac{1}{2} \|w\|_2^2 + C\xi$$

s.t. $\forall z \in \mathcal{Z}_\beta, \pi \in \Pi_{m, j_\beta} :$

$$w^\top (\phi_z(S, \pi^*) - \phi_z(S, \pi)) \geq \Delta_\beta(\pi^*, \pi) - \xi$$

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



SVMpAUC-tight: Cutting-Plane Solver

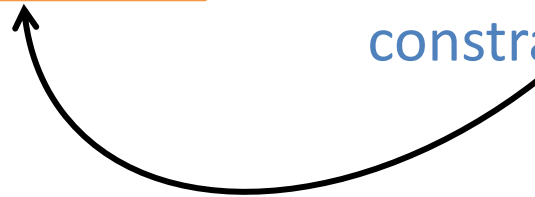
Repeat:

$$\min_{w, \xi \geq 0} \frac{1}{2} \|w\|_2^2 + C\xi$$

s.t. $\forall z \in \mathcal{Z}_\beta, \pi \in \Pi_{m, j_\beta} :$

$$w^\top (\phi_z(S, \pi^*) - \phi_z(S, \pi)) \geq \Delta_\beta(\pi^*, \pi) - \xi$$

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



Better Runtime Guarantees:
Maximum number of iterations
Time taken per iteration

SVMpAUC-tight: Projected Subgradient Solver

Primal formulation:

$$\min_w \left[\max_{z \in \mathcal{Z}_\beta, \pi \in \Pi_{m, j_\beta}} \Delta_\beta(\pi^*, \pi) - w^\top (\phi_z(S, \pi^*) - \phi_z(S, \pi)) \right]$$

s.t.

$$\|w\|_2 \leq \lambda$$

SVMpAUC-tight: Projected Subgradient Solver

Primal formulation:

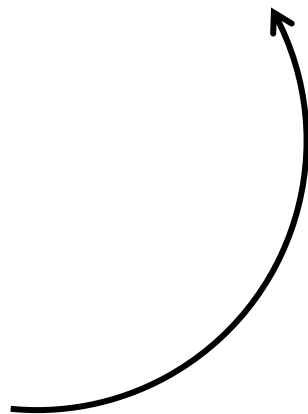
$$\min_w \left[\max_{z \in \mathcal{Z}_\beta, \pi \in \Pi_{m, j_\beta}} \Delta_\beta(\pi^*, \pi) - w^\top (\phi_z(S, \pi^*) - \phi_z(S, \pi)) \right]$$

s.t.

$\|w\|_2 \leq \lambda$

Repeat:

1. Compute **subgradient** and perform update
2. **Project on to the constraint set.**



SVMpAUC-tight: Projected Subgradient Solver

Primal formulation:

$$\min_w \left[\max_{z \in \mathcal{Z}_\beta, \pi \in \Pi_{m, j_\beta}} \Delta_\beta(\pi^*, \pi) - w^\top (\phi_z(S, \pi^*) - \phi_z(S, \pi)) \right]$$

s.t.

$$\|w\|_2 \leq \lambda$$

Sparsity-inducing
regularizations

LASSO
Group LASSO
Elastic-Net

Repeat:

1. Compute **subgradient** and perform update
2. **Project on to the constraint set.**

Outline

- Overview of SVMpAUC
- Upper Bound Optimized by SVMpAUC
- Improved Formulation: SVMpAUC-tight
- Optimization Methods
- Experiments

SVMpAUC-tight Vs SVMpAUC

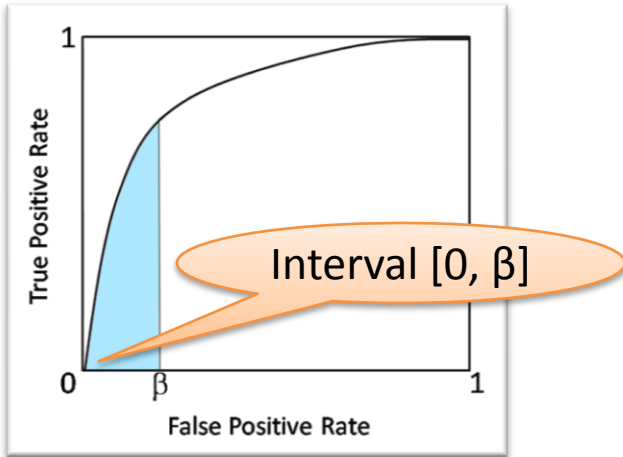
Partial AUC in [0, 0.1]

	Leukemia	PPI	Chem-informatics	KDD Cup 2001	Ovarian Cancer
SVMpAUC-tight	30.44	52.95	65.30	69.91	91.84
SVMpAUC	24.64	51.96	65.28	70.12	91.84
SVMAUC	28.83	39.72	62.78	62.23	92.17

Partial AUC in [0.2s, 0.3s]

	KDD Cup 2008
SVMpAUC-tight	53.43
SVMpAUC	51.89
SVMAUC	50.66

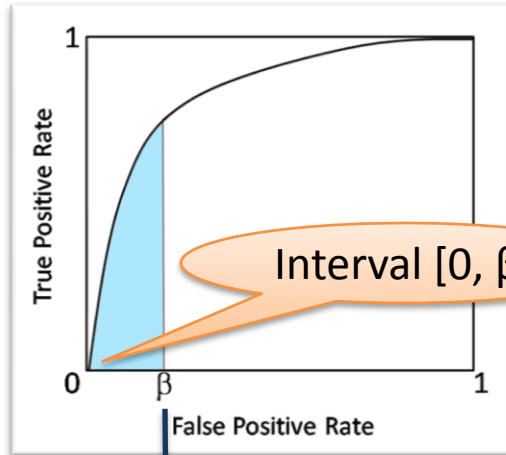
Run-time Analysis



Repeat:

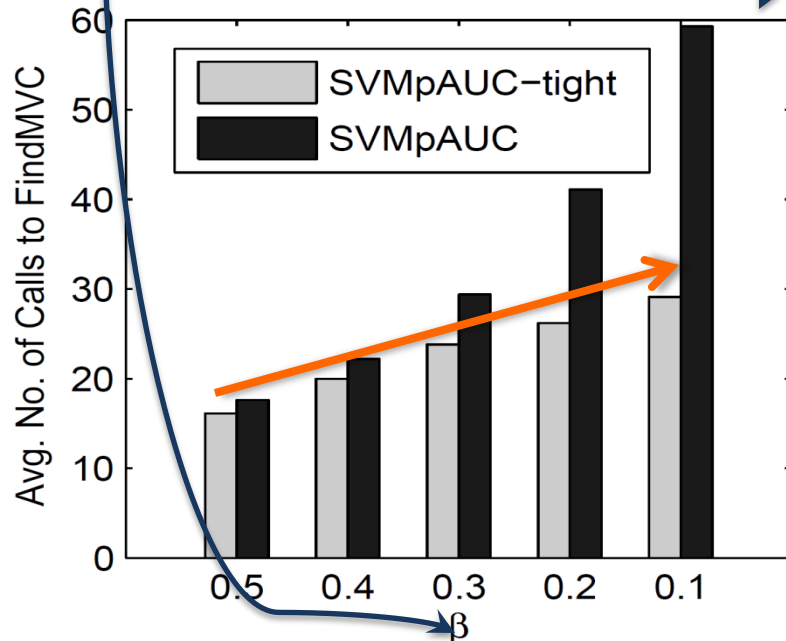
1. Solve OP for a subset of constraints.
2. Add the **most violated constraint**.

Run-time Analysis

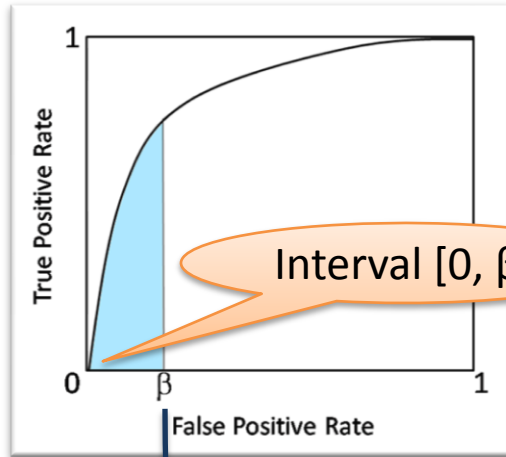


Repeat:

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

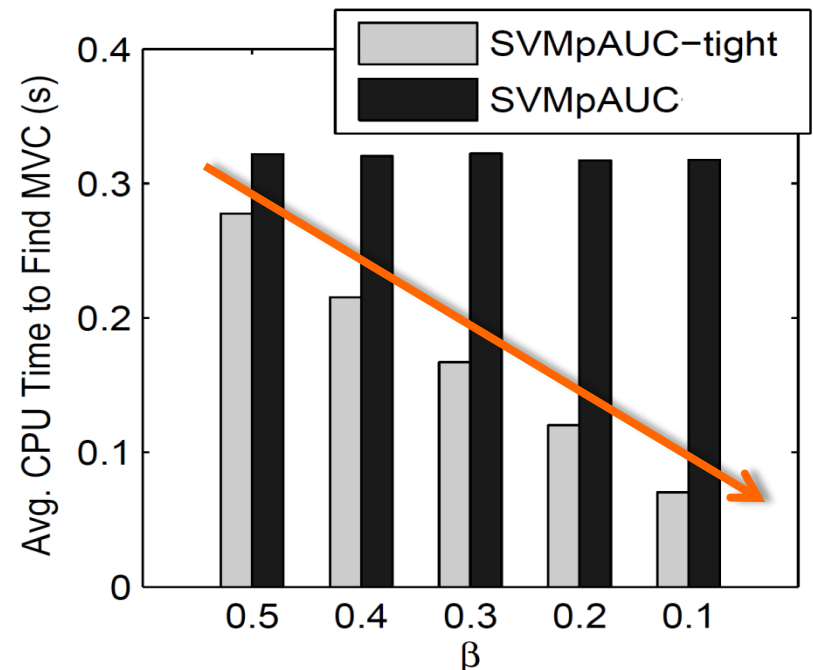
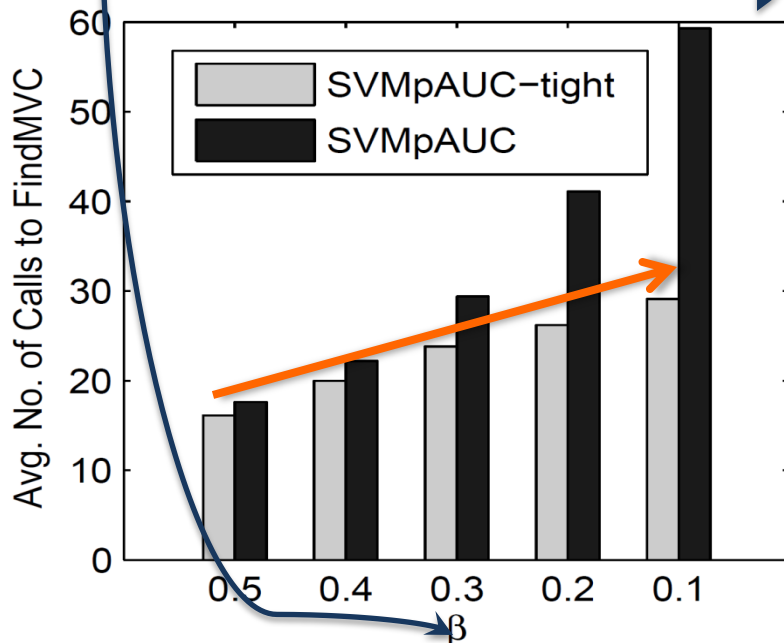


Run-time Analysis

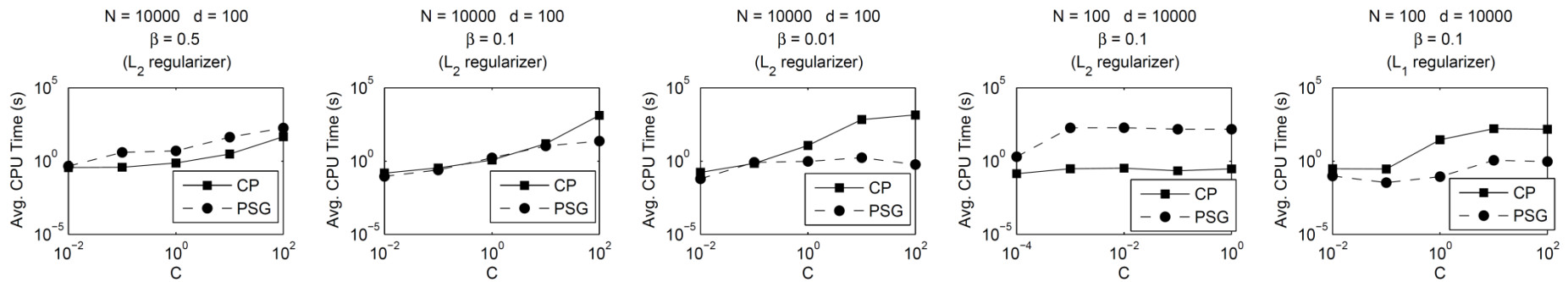


Repeat:

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



Cutting-Plane vs. Projected Subgradient



Cutting-plane method is faster on high dimensional data with L2 regularization

Projected subgradient method is faster with L1 regularization

Sparse and Group Sparse Extensions

	pAUC(0, 0.1)	
	Cheminformatics	KDD Cup 2001
$\text{SVM}_{\text{pAUC}}^{\ell_2}[0, 0.1]$	63.25 (100)	77.20 (100)
$\text{SVM}_{\text{pAUC}}^{\text{elastic-net}(0.001)}[0, 0.1]$	63.11 (41.5)	77.52 (41.6)
$\text{SVM}_{\text{pAUC}}^{\text{elastic-net}(0.1)}[0, 0.1]$	56.93 (32.24)	71.93 (27.6)
$\text{SVM}_{\text{pAUC}}^{\ell_1}[0, 0.1]$	53.63 (11.36)	66.22 (10.0)

	pAUC(0, 0.1)	# of groups selected
$\text{SVM}_{\text{pAUC}}^{\ell_2}[0, 0.1]$	67.09	17
$\text{SVM}_{\text{pAUC}}^{\ell_1/\ell_2}[0, 0.1]$	65.67	11.3

Sparse models at the cost of decrease in accuracy

Conclusions

- A new support vector algorithm for optimizing partial AUC based on a tight convex upper bound
- Cutting-plane solver with better run-time guarantees
- Experiments on several bioinformatics tasks demonstrate improved accuracy
- Projected subgradient solver allows sparse and group sparse extensions

Questions?