SUPPORT VECTOR MACHINES FOR DIFFERENTIAL PREDICTION

Finn Kuusisto[1], Vitor Santos Costa[2], Houssam Nassif[3], Elizabeth Burnside[1], David Page[1], and Jude Shavlik[1]

[1]University of Wisconsin – Madison [2]University of Porto [3]Amazon

DIFFERENTIAL PREDICTION

Goal
Use modeling techniques to **gain insight** about the **differences between two subgroups** of a population.

UPLIFT MODELING
(RADCLIFFE & SIMPSON, 2008)

How do we choose which customers to target with some marketing activity?

UPLIFT MODELING

(RADCLIFFE & SIMPSON, 2008)

How do we choose which customers to target with some marketing activity?

Persuadables	Customers who respond positively to marketing activity.
Sure Things	Customers who respond positively regardless.
Lost Causes	Customers who respond negatively regardless.
Sleeping Dogs	Customers who respond negatively to marketing activity.

UPLIFT MODELING
(RADCLIFFE & SIMPSON, 2008)

True customer groups are unknown.

Target		Control	
Response	**No Response**	**Response**	**No Response**
Persuadables, Sure Things	Sleeping Dogs, Lost Causes	Sleeping Dogs, Sure Things	Persuadables, Lost Causes

UPLIFT MODELING

Lift

The number of **true positives** that a classifier achieves **at a given proportion of the population labeled positive.**

Uplift

The **difference in lift** produced by a classifier between target and control subgroups.

$$\boldsymbol{AUU = AUL_T - AUL_C}$$

TASK: ADVERSE COX-2 INHIBITOR EFFECTS

- Non-steroidal anti-inflammatory drug (NSAID)
- Significantly reduced occurrence of adverse gastrointestinal effects common to other NSAIDs (e.g. ibuprofen)
- Rapid and widespread acceptance for treatment of ailments such as arthritis
- Later clinical trials showed increased risk of myocardial infarction (MI), or “heart attack”

Identify patients who are susceptible to an increased risk of MI as a direct result of taking COX-2 inhibitors.

UPLIFT MODELING TO MEDICINE: COX-2 INHIBITORS

Want

Identify patients who demonstrate an increased risk of MI as a direct result of being treated with COX-2 inhibitors.

Main Assumption

Patients with an increased risk of MI due to treatment with COX-2 inhibitors are directly analogous to customers with an increased chance of buying due to targeting – the persuadables.

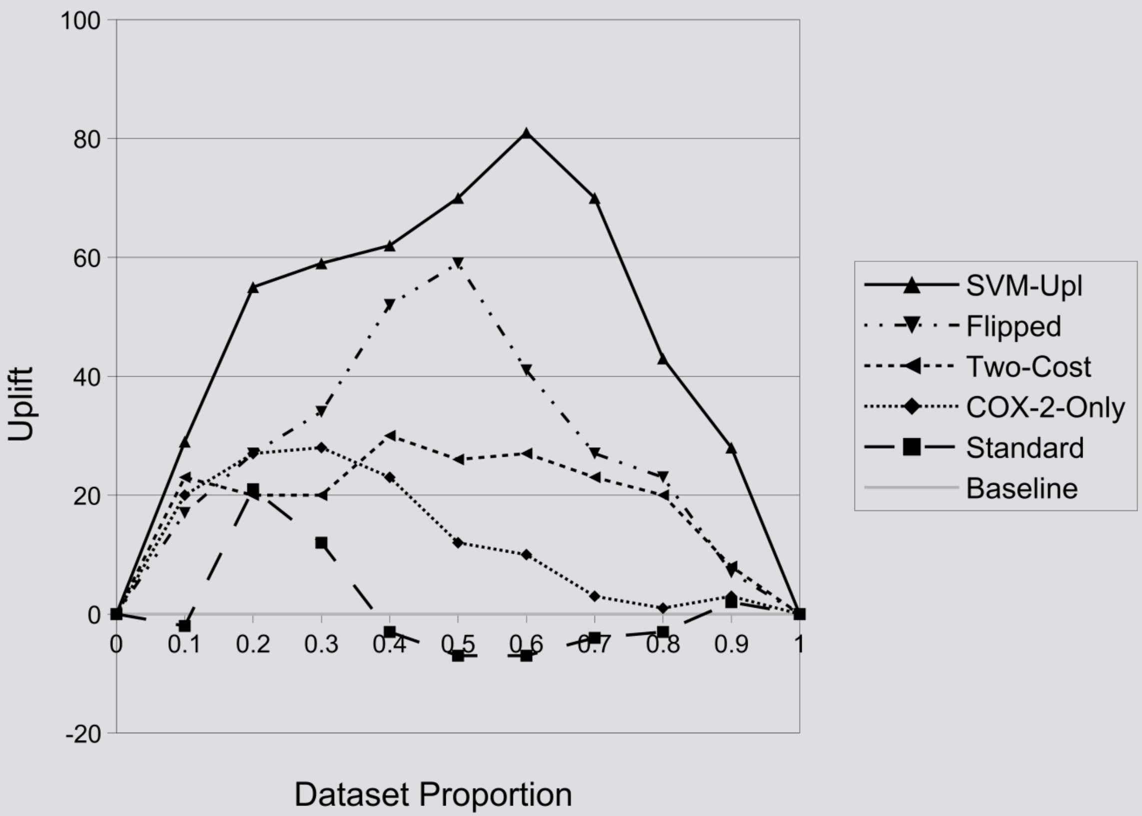
METHODS

- Compared SVM^{Upl} against 4 alternate SVM methods
- 10-fold cross-validation for evaluation
- Cost parameters selected from 10 through 10^{-6}
- Mann-Whitney test at 95% confidence for per-fold AUU comparison

RESULTS: COX-2 INHIBITORS

Model	**AUU**	**COX-2 AUL**	**No COX-2 AUL**	SVM^{Upl} ***p*-value**	
SVM^{Upl}	50.7	123.4	72.7	-	
Two-Cost	20.0	126.2	106.3	**0.004**	*
COX-2-Only	13.8	151.5	137.7	**0.002**	*
Standard	1.2	147.7	146.5	**0.002**	*
Flipped	28.5	102.2	73.6	**0.037**	*
Baseline	0.0	0.0	0.0	**0.002**	*

RESULTS: COX-2 INHIBITORS



HOW

Extend previous SVM work maximizing AUC (Joachims, 2005) to maximize AUU instead.

SVM FOR UPLIFT

Let the positive skew of data be:

$$\pi = \frac{P}{P + N}$$

Then (Tuffery, 2011):

$$AUL = P \times \left(\frac{\pi}{2} + (1 - \pi)AUC\right)$$

SVM FOR UPLIFT

$$AUU = AUL_T - AUL_C = P_T \times \left(\frac{\pi_T}{2} + (1-\pi_T)AUC_T\right) - P_C \times \left(\frac{\pi_C}{2} + (1-\pi_C)AUC_C\right)$$

$$max(AUU) \equiv max(P_T \times (1-\pi_T)AUC_T - P_C \times (1-\pi_C)AUC_C)$$

$$\propto max\left(AUC_T - \underbrace{\frac{P_C \times (1-\pi_C)}{P_T \times (1-\pi_T)}}_{\lambda} AUC_C\right)$$

$$max(AUU) \equiv max(AUC_T - \lambda AUC_C)$$

TASK: IN SITU BREAST CANCER

- Most common cancer in women
- Two basic stages: *In situ* and *invasive*
  - *In situ* cancer cells are localized
  - *Invasive* cancer cells have infiltrated surrounding tissue
- Younger women tend to have more aggressive in situ cancer
- Older women sometimes have indolent in situ cancer

Identify older patients with indolent in situ breast cancer.

UPLIFT MODELING TO MEDICINE: BREAST CANCER

Want

Identify older patients with in situ breast cancer that is distinct from that of younger patients.

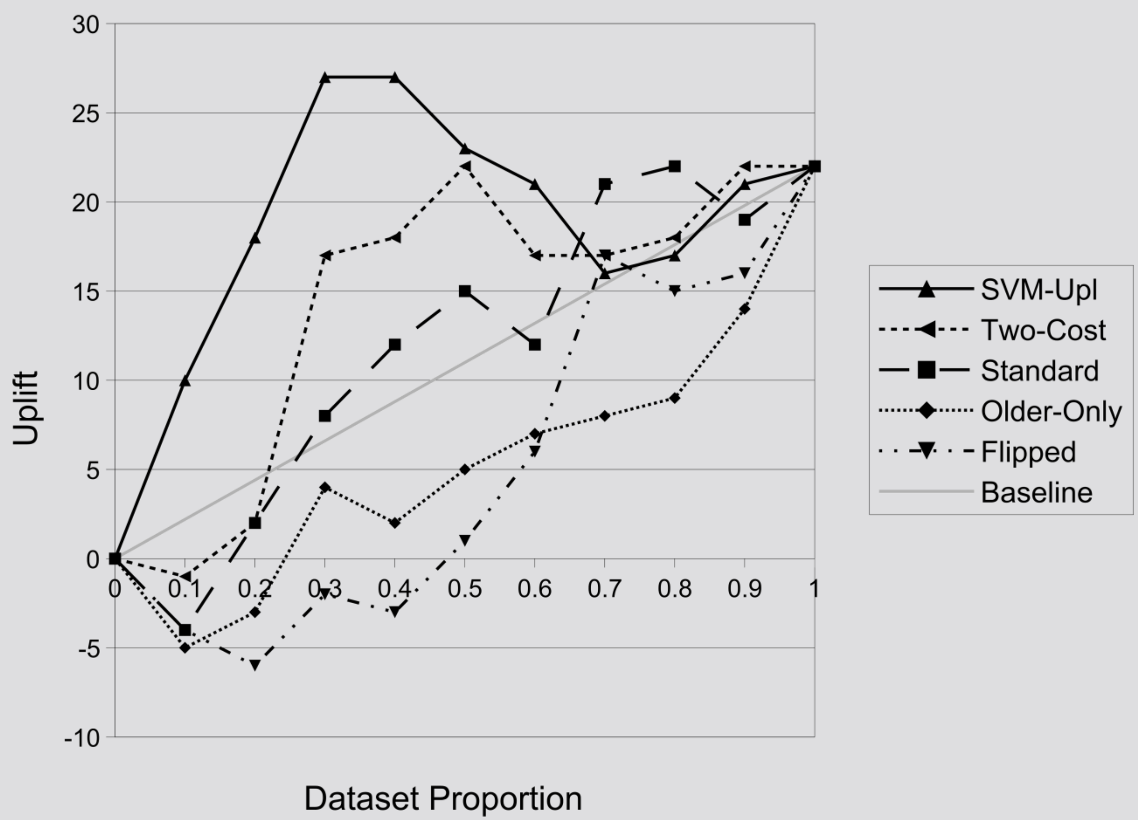
Main Assumption

Older patients with in situ breast cancer that is distinct from that of younger patients, who tend to have aggressive cancer, have a decreased risk of invasive progression.

RESULTS: BREAST CANCER

Model	**AUU**	**Older AUL**	**Younger AUL**	***SVM^{Upl}* *p*-value**	
SVM^{Upl}	19.2	64.3	45.1	-	
Two-Cost	13.5	74.3	60.8	0.432	
Older-Only	5.9	67.7	61.9	**0.037**	*
Standard	11.0	75.4	64.3	**0.049**	*
Flipped	4.8	53.9	49.1	**0.020**	*
Baseline	11.0	66.0	55.0	**0.004**	*

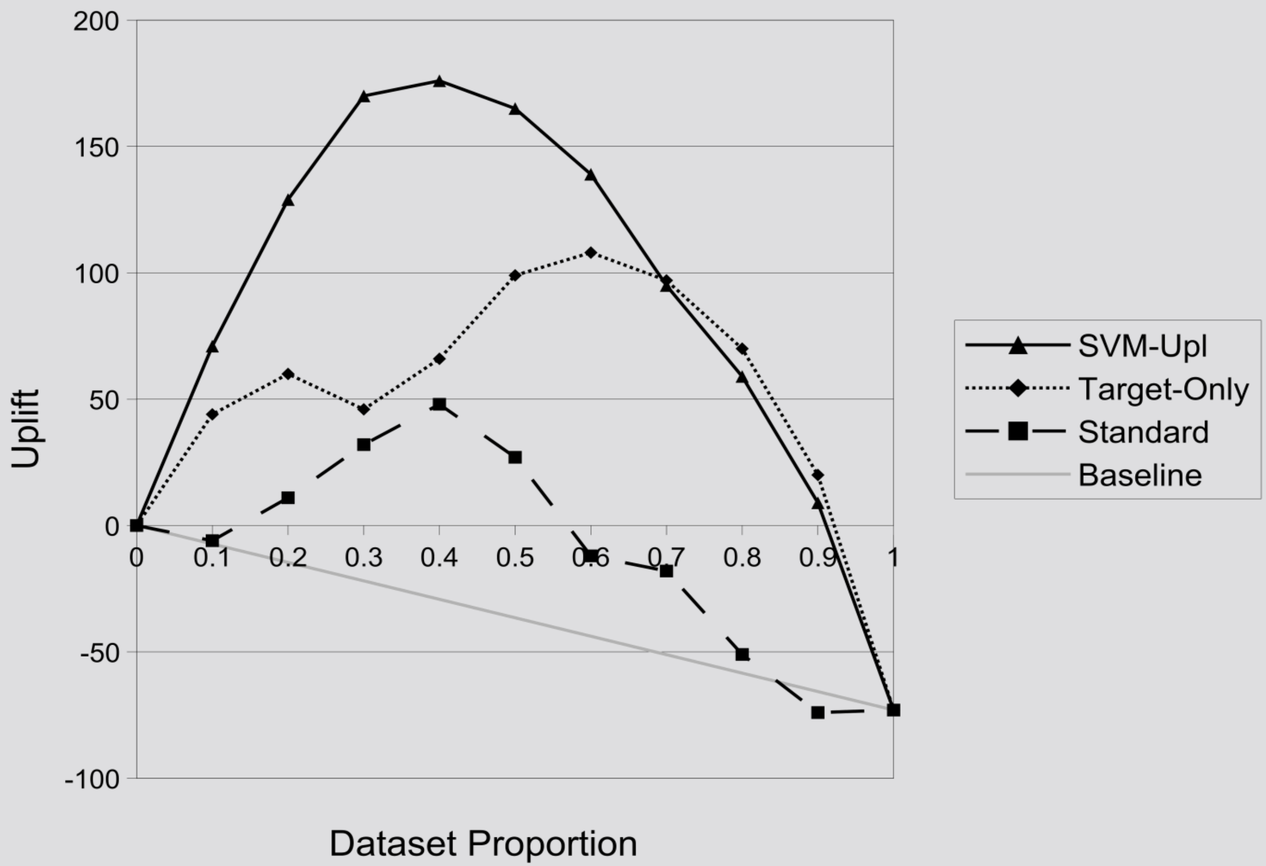
RESULTS: BREAST CANCER



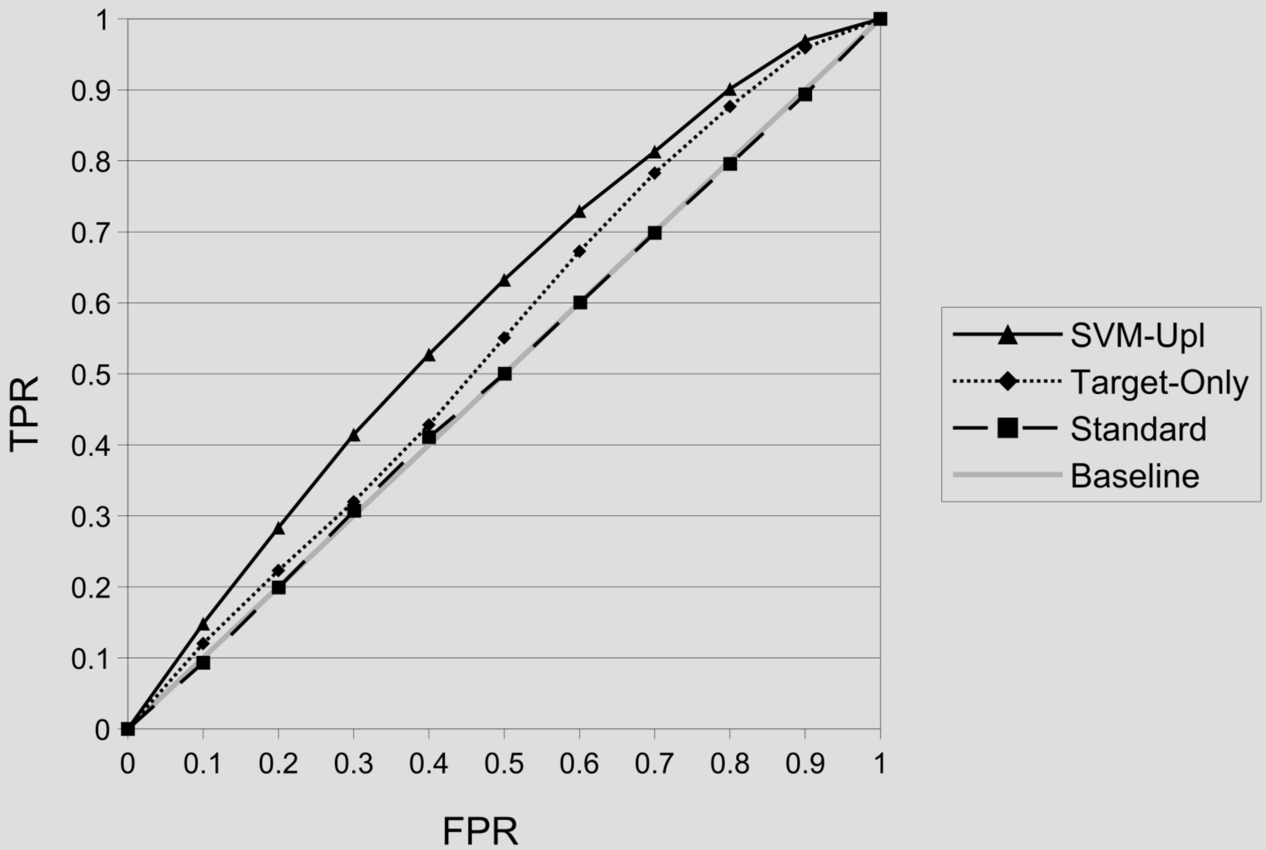
UPLIFT MODELING SIMULATION

- Generated synthetic customer population
- Subjected customer population randomly to simulated marketing activity
- Measured uplift as usual
- Measured ROC with *Persuadables* as the positive class, others as negative

UPLIFT MODELING SIMULATION: UPLIFT CURVE



UPLIFT MODELING SIMULATION: PERSUADABLE ROC



CONCLUSIONS & FUTURE WORK

- Extended previous SVM work on AUC maximization to AUU
- Results suggest SVM^{Upl} achieves better uplift than many alternate SVM methods

- May want to make performance guarantees for control group
- May want to interpret learned model
- Better verification that maximizing uplift is appropriate goal

THANKS

Questions?

SELECTED REFERENCES

Radcliffe, N. and Simpson. R.: Identifying who can be saved and who will be driven away by retention activity. Journal of Telecommunications Management (2008).

Tuffery, S.: Data Mining and Statistics for Decision Making. John Wiley & Sons, 2nd edn. (2011).

Joachims, T.: A support vector method for multivariate performance measuers. In: Proceedings of the 22nd International Conference on Machine Learning (2005).

APPENDIX

SVM FOR ROC

(JOACHIMS, 2005)

Define y'_{ij} as a predicted label on pairs of positive (i) and negative (j) examples, where $y'_{ij} = 1$ if $y'_i > y'_j$, and -1 otherwise.

Joachims' algorithm to maximize AUC corresponds to finding y^*_{ij} that maximizes:

$$\mathbf{w}^T \Psi(\bar{x}, \bar{y}') + \Delta_{AUC}(\bar{y}', \bar{y})$$

Where:

$$\mathbf{w}^T \Psi(\bar{\mathbf{x}}, \bar{y}') = \mathbf{w}^T \frac{1}{2} \sum_{i=1}^{P} \sum_{j=1}^{N} y'_{ij} \left(\mathbf{x}_i - \mathbf{x}_j \right) \qquad \Delta_{AUC}(\bar{y}', \bar{y}) = \sum_{i=1}^{P} \sum_{j=1}^{N} \frac{1}{2} \left(1 - y'_{ij} \right)$$

SVM FOR UPLIFT

Let the positive skew of data be:

$$\pi = \frac{P}{P + N}$$

Then:

$$AUL = P\left(\frac{\pi}{2} + (1 - \pi)AUC\right)$$

SVM FOR UPLIFT

$$AUU = P_A\left(\frac{\pi_A}{2} + (1-\pi_A)AUC_A\right) - P_B\left(\frac{\pi_B}{2} + (1-\pi_B)AUC_B\right)$$

$$max(AUU) \equiv max(P_A(1-\pi_A)AUC_A - P_B(1-\pi_B)AUC_B)$$
$$\propto max\left(AUC_A - \frac{P_B(1-\pi_B)}{P_A(1-\pi_A)}AUC_B\right)$$

Defining:

$$\lambda = \frac{P_B(1-\pi_B)}{P_A(1-\pi_A)}$$

Then:

$$max(AUU) \equiv max(AUC_A - \lambda AUC_B)$$

SVM FOR UPLIFT

Now simply redefine the AUC optimization:

$$\mathbf{w}^T\Psi(\bar{\mathbf{x}},\bar{y}') = \mathbf{w}^T\frac{1}{2}\sum_{i=1}^{P_A}\sum_{j=1}^{N_A} y'_{ij}\left(\mathbf{x}_i - \mathbf{x}_j\right) + \lambda\mathbf{w}^T\frac{1}{2}\sum_{k=1}^{P_B}\sum_{l=1}^{N_B} y'_{lk}(\mathbf{x}_l - \mathbf{x}_k)$$

$$\Delta_{AUU}(\bar{y}',\bar{y}) = \sum_{i=1}^{P_A}\sum_{j=1}^{N_A}\frac{1}{2}\left(1 - y'_{ij}\right) + \lambda\sum_{k=1}^{P_B}\sum_{l=1}^{N_B}\frac{1}{2}(1 - y'_{lk})$$

UPLIFT MODELING SIMULATION

1. Generated random Bayesian network with 20 nodes and 30 edges
2. Selected one four-value node to define the customer group
3. Drew 10,000 samples from the network
4. Subjected customer population randomly to simulated marketing activity
   - *Persuadables* always responded when targeted, never when not
   - *Sleeping Dogs* never responded when targeted, always when not
   - *Sure Things* and *Lost Causes* always and never responded respectively

- Measured uplift as usual
- Measured ROC with *Persuadables* as the positive class, others as negative

MATERIALS

COX-2 Inhibitors		No COX-2 Inhibitors	
MI	No MI	MI	No MI
184	1,776	184	1,776

Older		Younger	
In Situ	Invasive	In Situ	Invasive
132	401	110	264