

# Seeing stars:

Exploiting class relationships for sentiment categorization w.r.t. rating scales

**Written by Bo Pang & Lillian Lee**



Presented by Katherine Schinkel & Marcus Rosti

# Sentiment Analysis Text Categorization

## Binary Classification



Good



Bad

**vs.**

## Multiclass Classification



Bad

Below Average

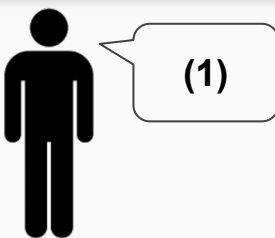
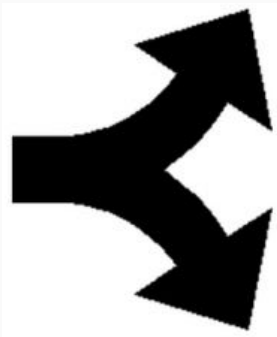
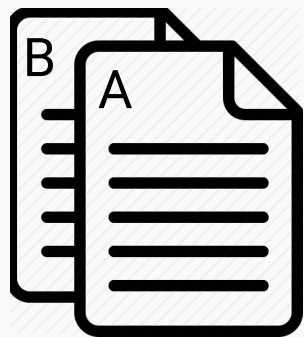
Average

Above Average

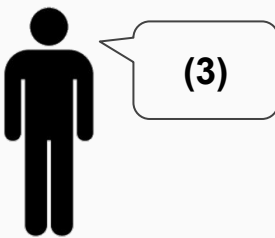
Good

# Can a human tell the difference?

Movie review pairs  
from same author



(1)



(3)

**Determine:**

(1) A is more positive than B

**OR**

(2) A is less positive than B

**OR**

(3) A is as positive as B

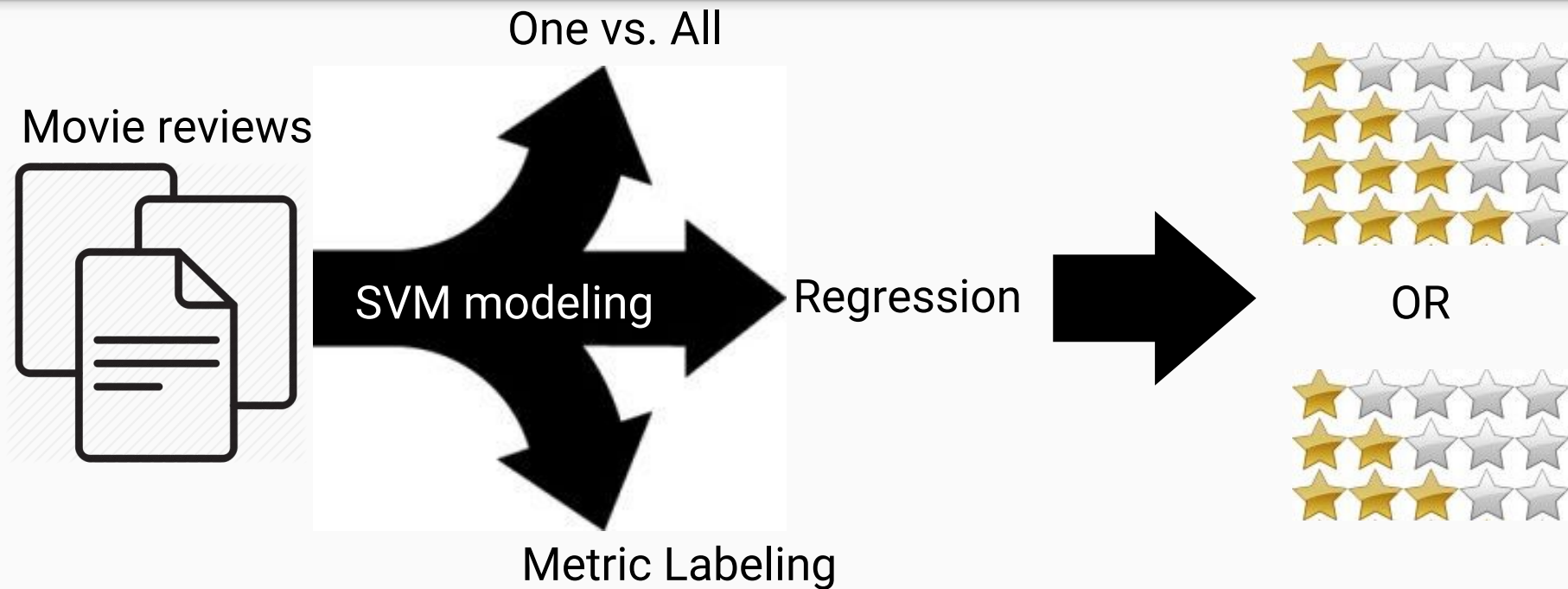
# Can a human tell the difference?

Rating diff.	Pooled	Subject 1	Subject 2
3 or more	100%	100% (35)	100% (15)
2 (e.g., 1 star)	83%	77% (30)	100% (11)
1 (e.g., $\frac{1}{2}$ star)	69%	65% (57)	90% (10)
0	55%	47% (15)	80% ( 5)

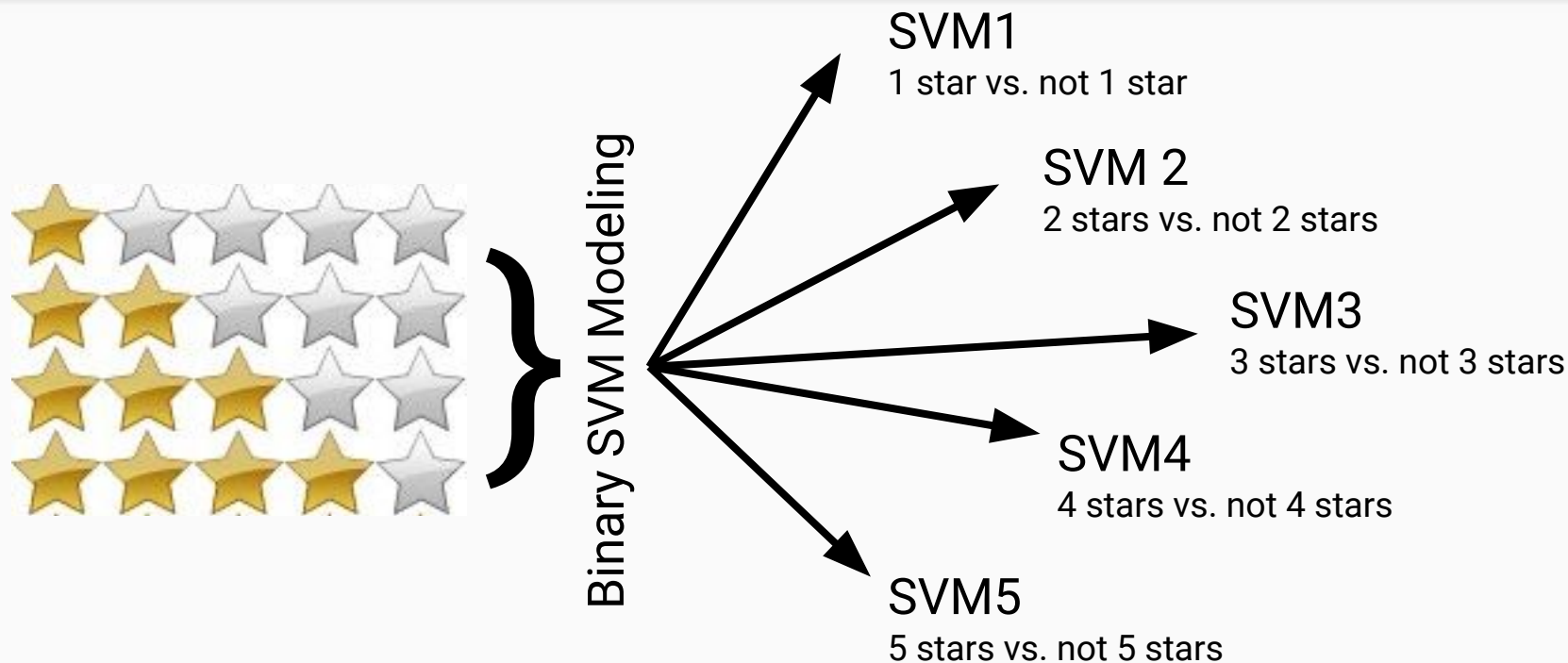
# Can a computer tell the difference?



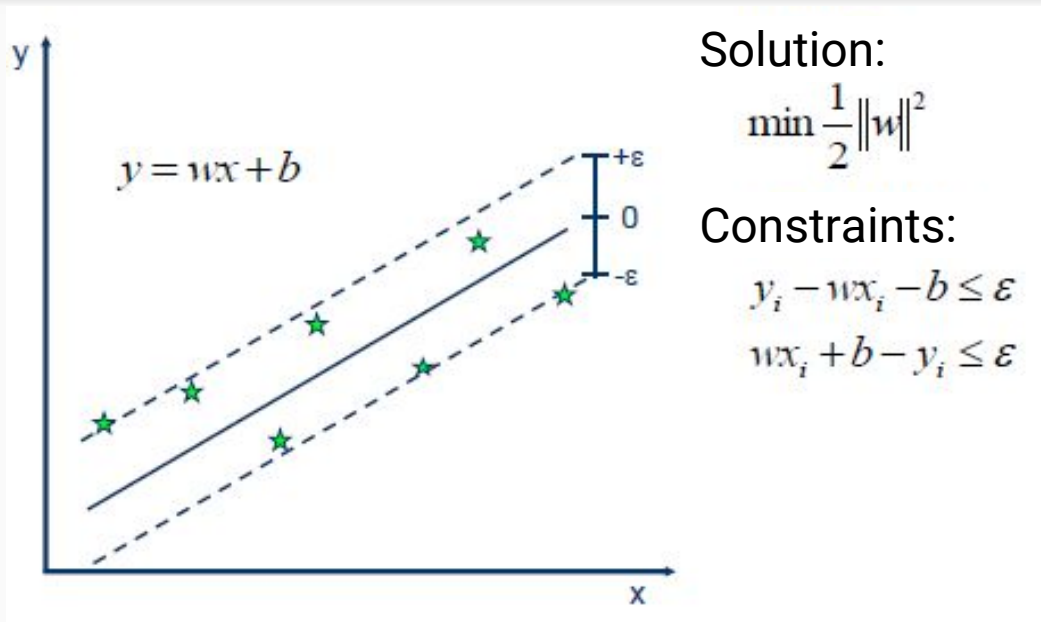
# Computer implementation



# Using One vs. All



# Using SVM regression



***implicitly*** encodes the “similar items, similar labels” heuristic



# Using Metric Labeling



is more similar to



Distance of 2 

than to



Distance of 6 

**explicitly** encodes the “similar items, similar labels” heuristic

# Feature Engineering: Positive Sentence Percentage (PSP)

$$\cos \left( \text{PSP}(x), \text{PSP}(y) \right) .$$



$$\frac{\# \text{ positive sentences}}{\# \text{ negative sentences}}$$



Sentence polarity dataset:



10,622 review snippets



One sentence long



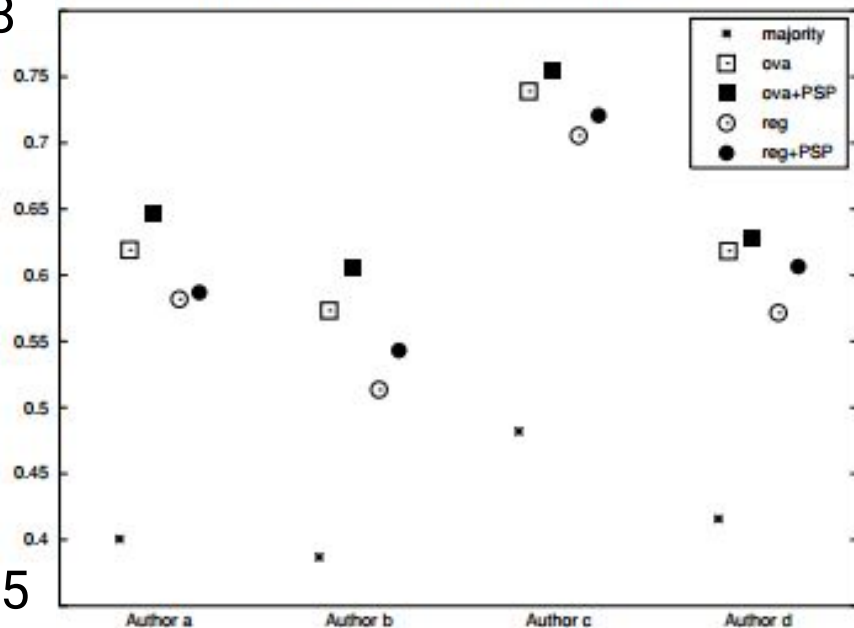
Labeled with review source  
label ("positive" or  
"negative")



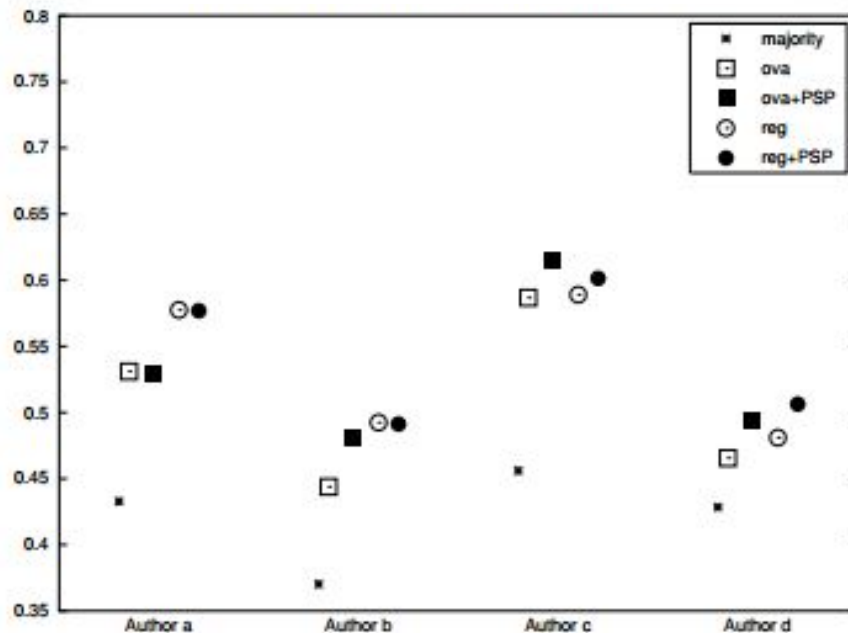
Naive Bayes Classifier

# Modeling Results

Average Accuracies, three-class data



Average Accuracies, four-class data



# Model Comparison

Significant differences, three-class data

	ova a b c d	ova+PSP a b c d	reg a b c d	reg+PSP a b c d
ova		△△△·	◁◁◁◁	·◁··
ova+PSP	◀◀◀·		◁◁◁◁	◁◁◁·
reg	△△△△	△△△△		·△·△
reg+PSP	·△··	△△△·	·◀·◀	

Significant differences, four-class data

	ova a b c d	ova+PSP a b c d	reg a b c d	reg+PSP a b c d
ova		·△△△	△△··	△··△
ova+PSP	·◀◀◀		△···	△···
reg	◁◁··	◁···		····
reg+PSP	◁··◁	◁···	····	

# Takeaways

## Improvements

Validates the 'rotten tomatoes' method

Evaluates modeling techniques

Can be used in aggregate across many authors  
as a multiclass classifier

## Criticisms

Doesn't compare to binary classification

Only really adds one class

{Good, Meh, Bad}