

# INTRODUCTION TO DATA SCIENCE

This lecture is  
based on course by E. Fox and C. Guestrin, Univ of Washington

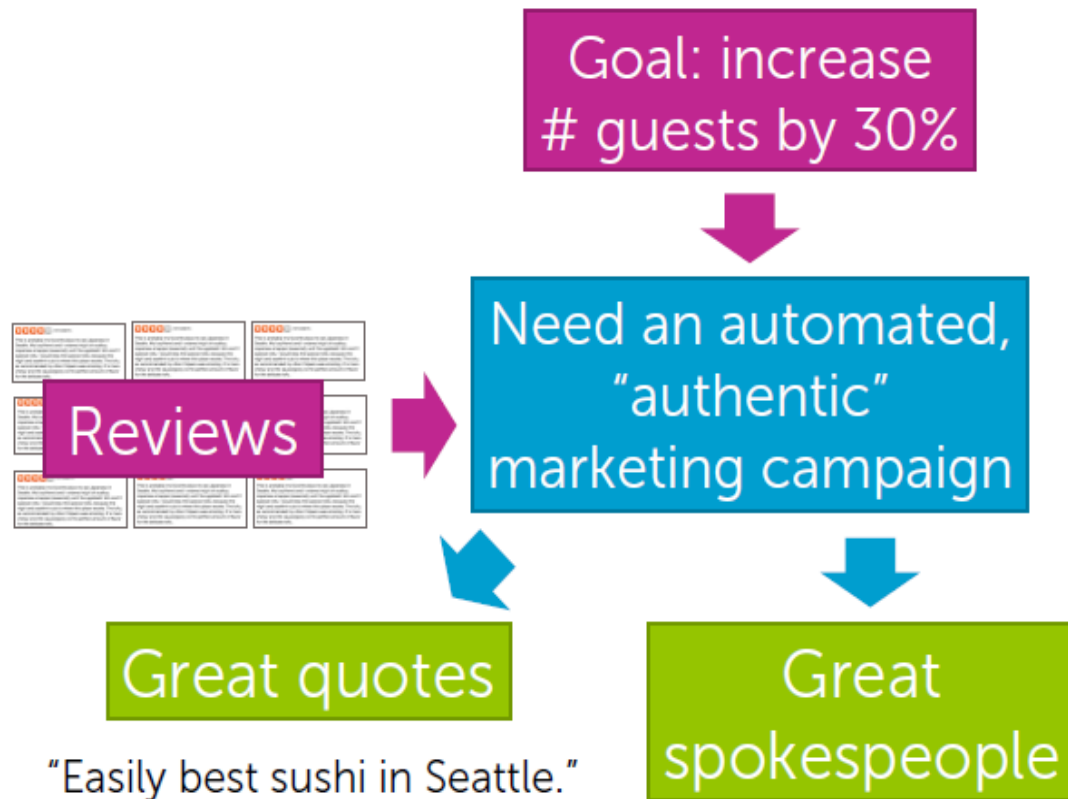
21/11/2017

WFAiS UJ, Informatyka Stosowana  
II stopień studiów

# Evaluating Classifiers

2

## Using reviews to promote my restaurant



# Classifiers

3

## How do I find sentences with positive sentiment?

All reviews  
for my restaurant

<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>	<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>	<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>
<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>	<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>	<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>
<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>	<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>	<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>
<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>	<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>	<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>
<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>	<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>	<p>★★★★ (10/10)</p> <p>This is a wonderful family restaurant. It has a friendly atmosphere and the food is excellent. The service is top-notch and the staff are very helpful. I highly recommend this restaurant to everyone.</p>

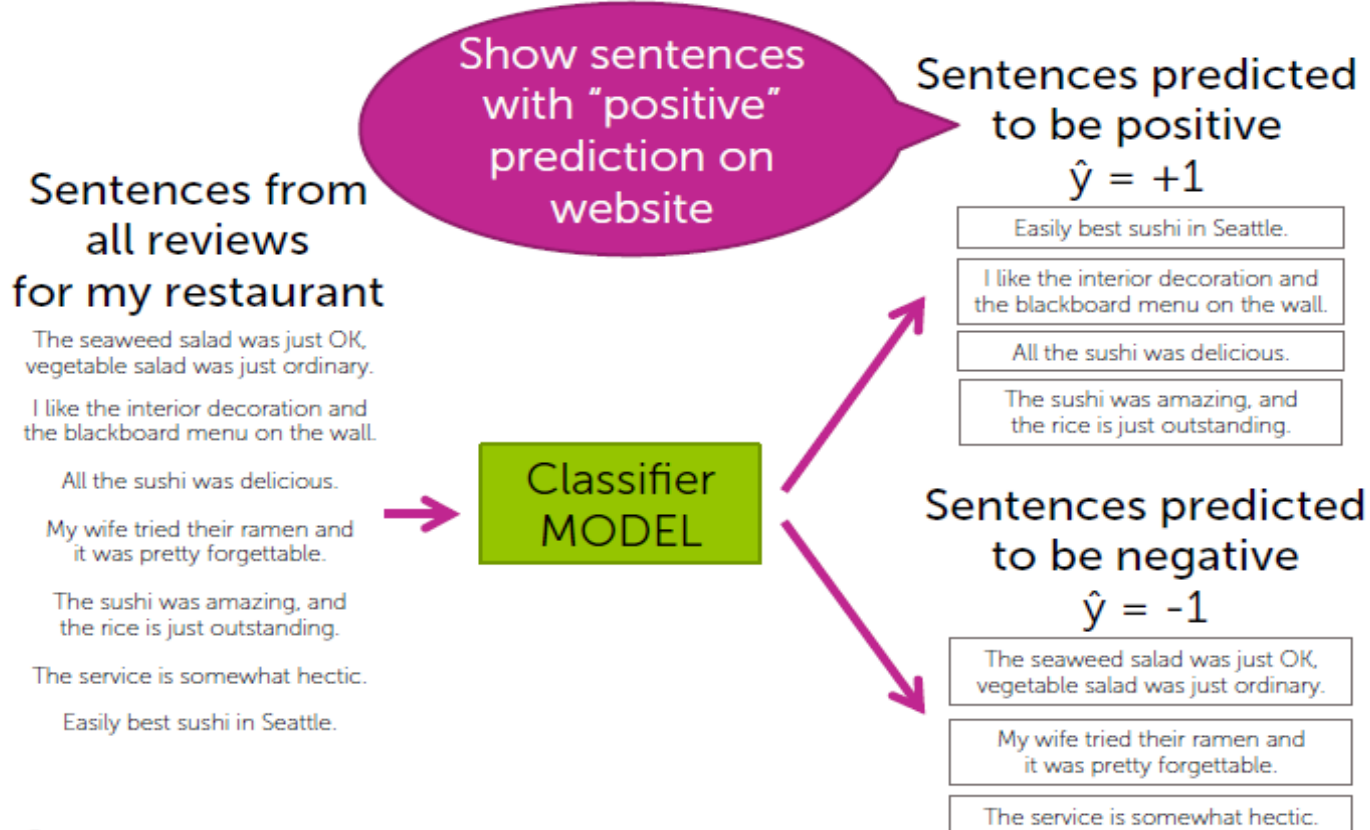
What are the positive things being said about my restaurant?



# Classifiers

4

## Use the sentiment classifier model!



# What does it mean for the classifiers to be good?

5

We explored accuracy of random classifier as baseline

- For binary classification:
  - Half the time, you'll get it right! (on average)
    - classification error = 0.5
- For  $k$  classes, error =  $1 - 1/k$ 
  - error = 0.666 for 3 classes, 0.75 for 4 classes,...

At the very, very, very least,  
you should healthily beat random...  
Otherwise, it's (usually) pointless...

# What does it mean for the classifiers to be good?

6

We explored the pitfalls of imbalanced problems:  
*Is 90% accuracy good? Depends ...*

*90% of sentences are negative!*

```
graph TD; A["90% of sentences are negative!"] --> B["90% accuracy by predicting every sentence is negative!!!"]; B --> C["Amazing 'performance' but not useful for me right now!"];
```

90% accuracy by predicting every sentence is negative!!!

Amazing "performance" but not useful for me right now!

# Precision & Recall

7

Automated marketing campaign cares about something else...

Website shows 10 sentences from recent reviews



**PRECISION**

Did I (mistakenly) show a negative sentence???



**RECALL**

Did I not show a (great) positive sentence???

Accuracy doesn't capture these issues well...

# Precision & Recall

8

## **Precision:**

Fraction of positive predictions that are actually positive

## **Recall:**

Fraction of positive data predicted to be positive



# Precision

9

What fraction of the positive predictions are correct?

Sentences predicted to be positive:  $\hat{y}_i = +1$

Easily best sushi in Seattle.	✓
The seaweed salad was just OK, vegetable salad was just ordinary.	✗
I like the interior decoration and the blackboard menu on the wall.	✓
The service is somewhat hectic.	✗
The sushi was amazing, and the rice is just outstanding.	✓
All the sushi was delicious.	✓

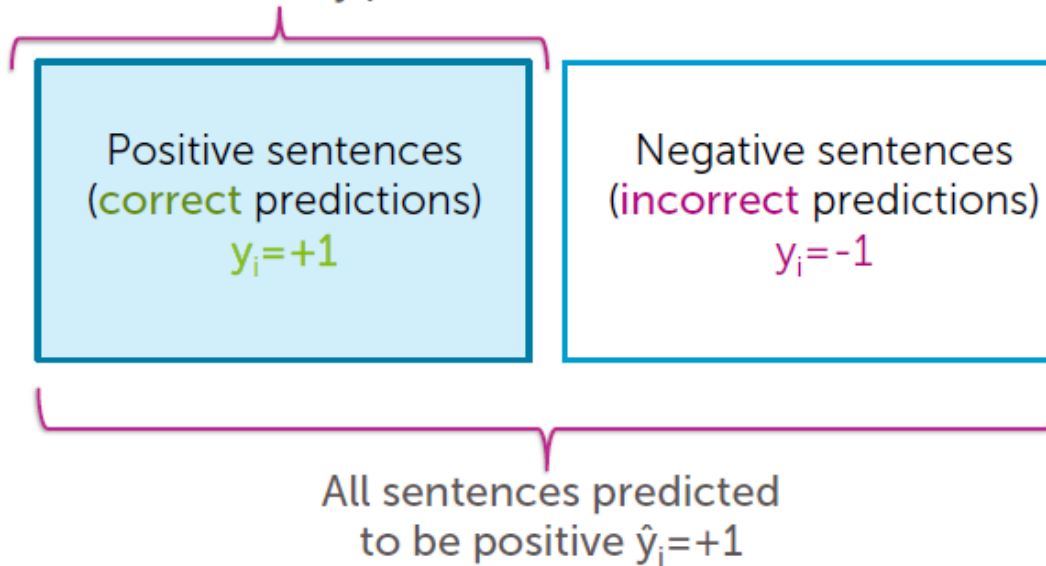
Only 4 out of 6 sentences predicted to be **positive** are actually **positive**

# Precision

10

**Precision:** Fraction of positive predictions that are actually positive





Subset of positive predictions that are actually positive



# Precision

11

## Types of error: *Review*

		Predicted label	
		 $\hat{y}_i = +1$	 $\hat{y}_i = -1$
True label	 $y_i = +1$	True Positive	False Negative
	 $y_i = -1$	False Positive	True Negative

# Precision

12

## Precision - Formula

- Fraction of positive predictions that are correct

$$\text{precision} = \frac{\# \text{ true positives}}{\# \text{ true positives} + \# \text{ false positives}}$$

- Best possible value : 1.0
- Worst possible value : 0.0

# Precision

13

## Why precision is important

Shown on website

Sentences predicted to be positive:  $\hat{y}_i = +1$

Easily best sushi in Seattle.	✓
The seaweed salad was just OK, vegetable salad was just ordinary.	✗
I like the interior decoration and the blackboard menu on the wall.	✓
The service is somewhat hectic.	✗
The sushi was amazing, and the rice is just outstanding.	✓
All the sushi was delicious.	✓

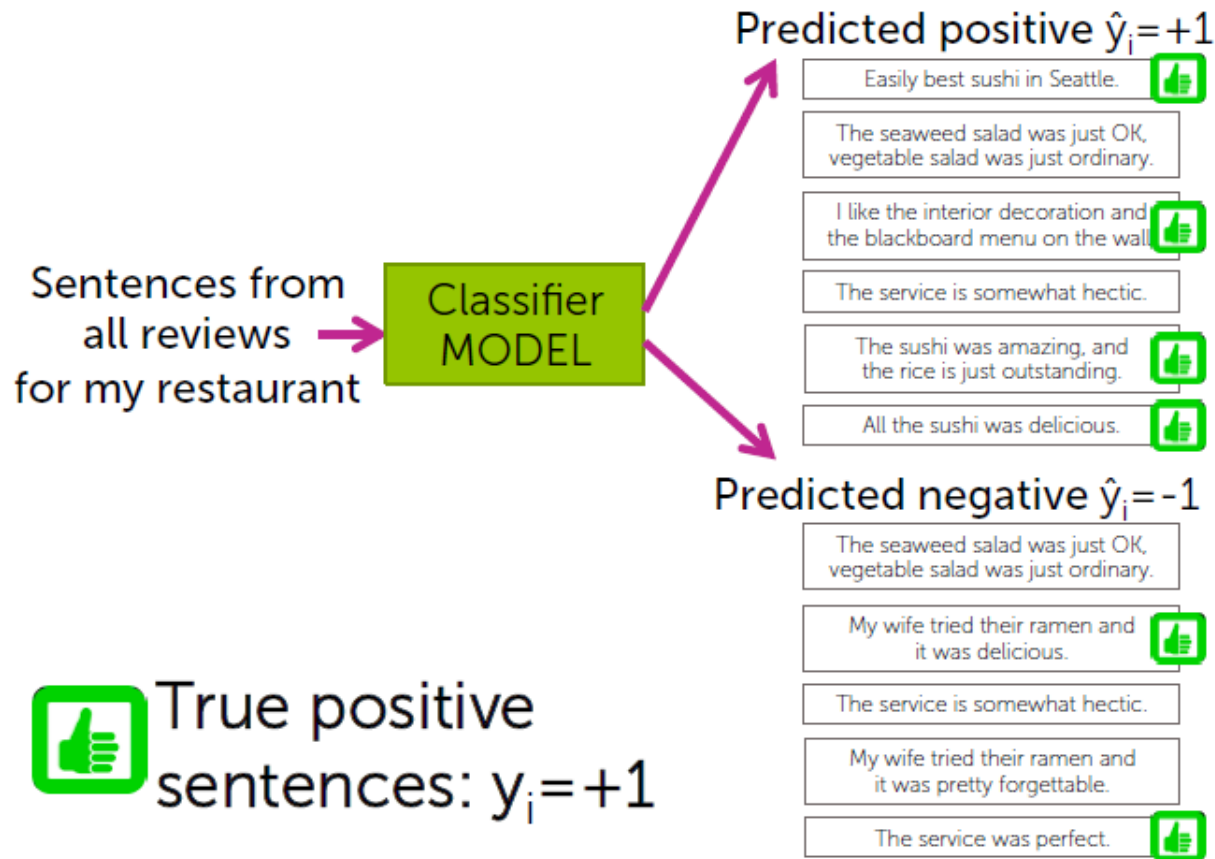
2 negative sentences shown to potential customers... ☹️

High precision means positive predictions actually likely to be positive!

# Recall

14

## Did I find all the positive sentences?







# Recall

15

What fraction of positive sentences were missed out?



**Predicted positive  $\hat{y}_i = +1$**

- Easily best sushi in Seattle. 
- The seaweed salad was just OK, vegetable salad was just ordinary.
- I like the interior decoration and the blackboard menu on the wall. 
- The service is somewhat hectic.
- The sushi was amazing, and the rice is just outstanding. 
- All the sushi was delicious. 

← Found 4 positive sentences

**Model could not find 2 sentences that were actually positive**

**Predicted negative  $\hat{y}_i = -1$**

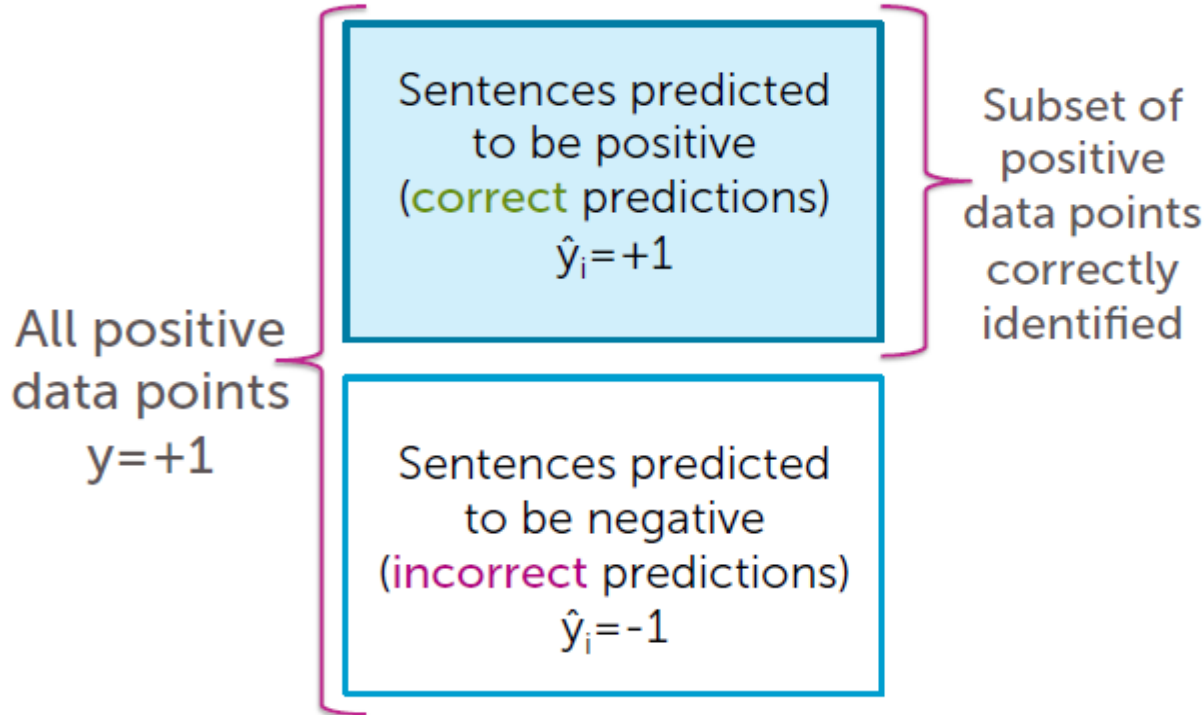
- The seaweed salad was just OK, vegetable salad was just ordinary.
- My wife tried their ramen and it was delicious. 
- The service is somewhat hectic.
- My wife tried their ramen and it was pretty forgettable.
- The service was perfect. 

← Missed 2 positive sentences

# Recall

16

Recall: Fraction of positive data predicted to be positive





# Recall

17

## Recall - Formula

- Fraction of positive data points correctly classified

$$\text{Recall} = \frac{\# \text{ true positives}}{\# \text{ true positives} + \# \text{ false negatives}}$$


- Best possible value : 1.0
- Worst possible value : 0.0

# Recall


18

## Why is recall important?


### Predicted positive $\hat{y}_i = +1$


Easily best sushi in Seattle. 

The seaweed salad was just OK,  
vegetable salad was just ordinary.

I like the interior decoration and  
the blackboard menu on the wall. 


The service is somewhat hectic.

The sushi was amazing, and  
the rice is just outstanding. 

All the sushi was delicious. 


### Predicted negative $\hat{y}_i = -1$

The seaweed salad was just OK,  
vegetable salad was just ordinary.


My wife tried their ramen and  
it was delicious. 

The service is somewhat hectic.

My wife tried their ramen and  
it was pretty forgettable.

The service was perfect. 

Want to show positive  
sentences on website

2 positive sentences  
not shown to potential  
customers... 

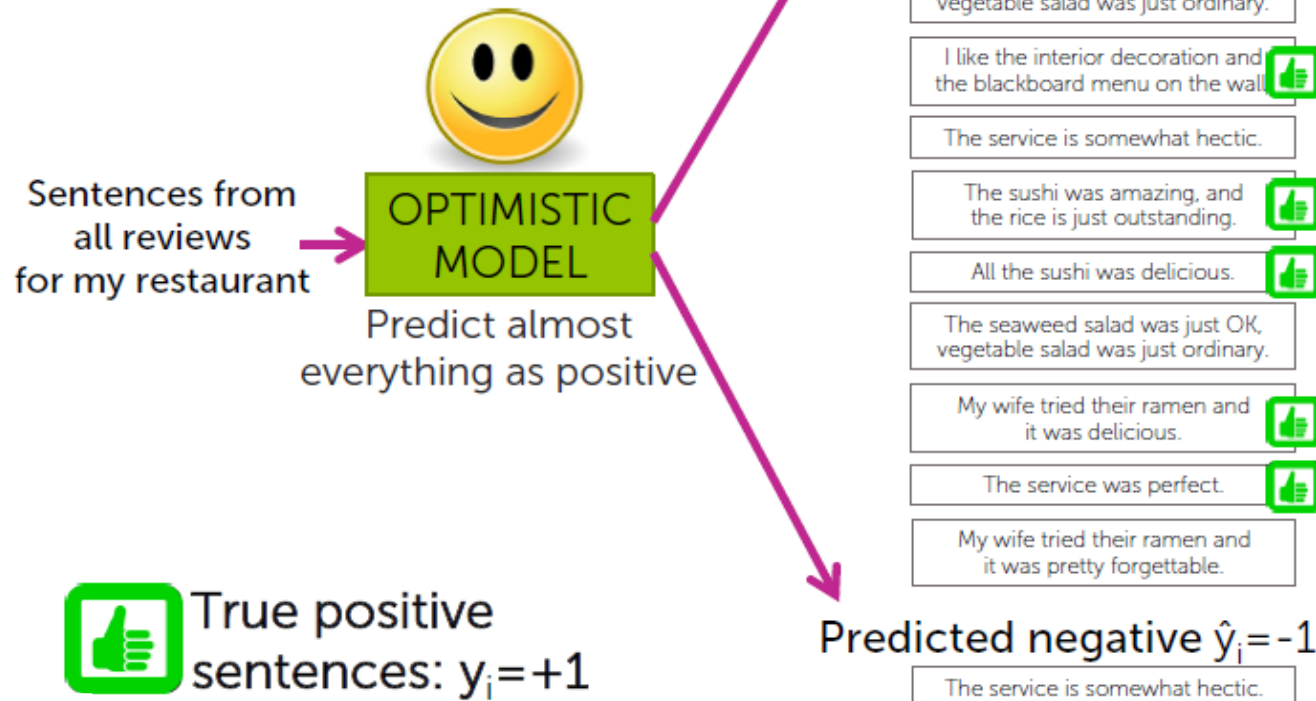
High recall  
means positive  
data points are  
very likely to be  
discovered!

# Precision-recall extremes

19

## Optimistic model:

High recall, low precision

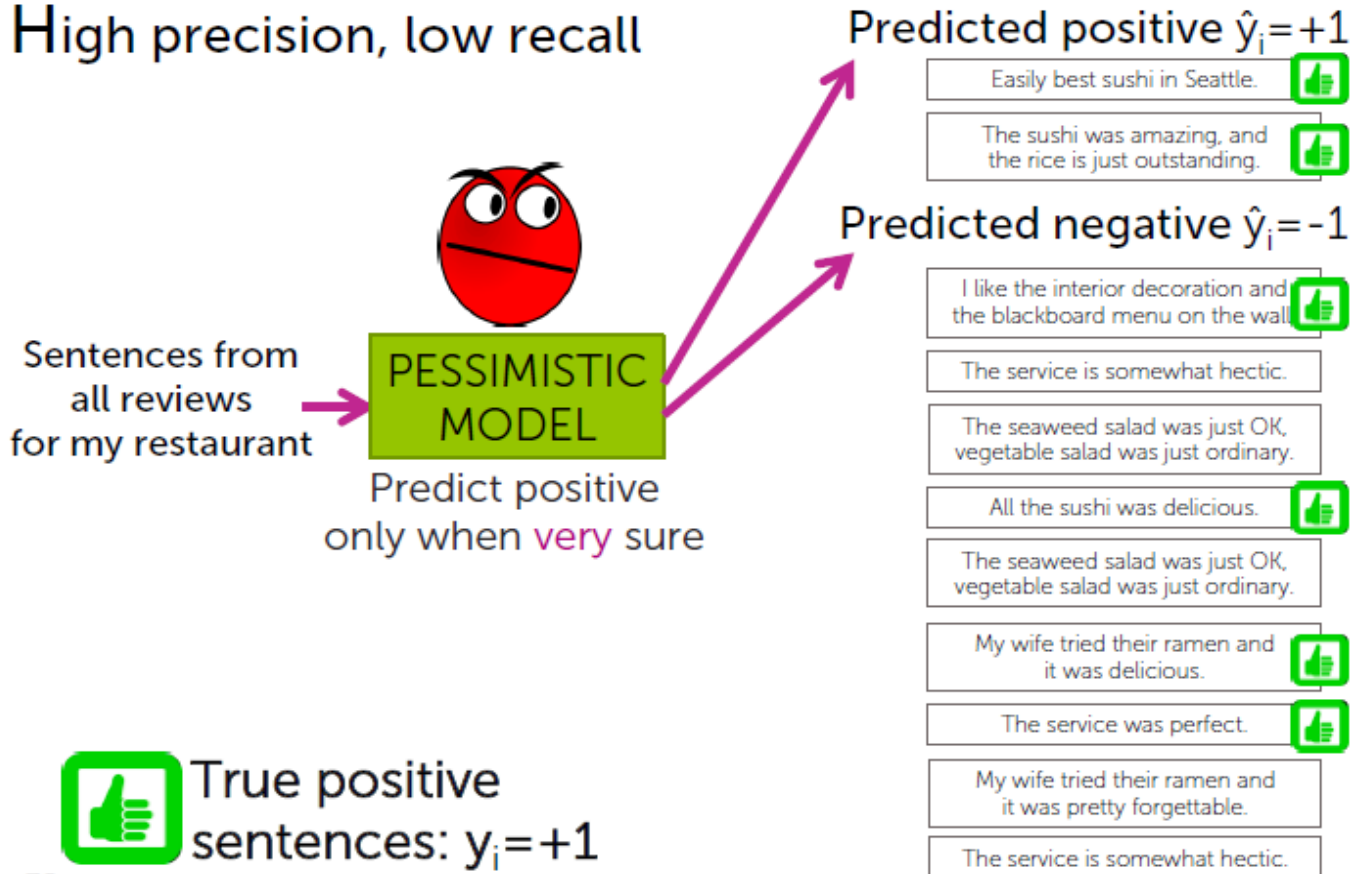


# Precision-recall extremes

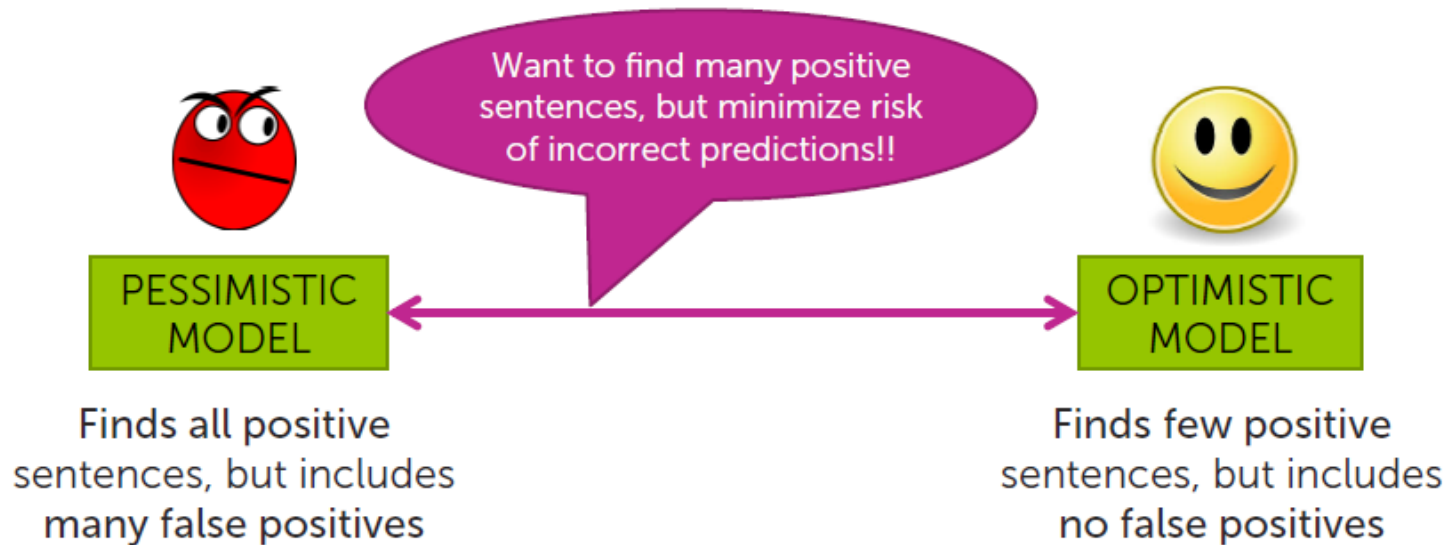
20

## Pessimistic model:

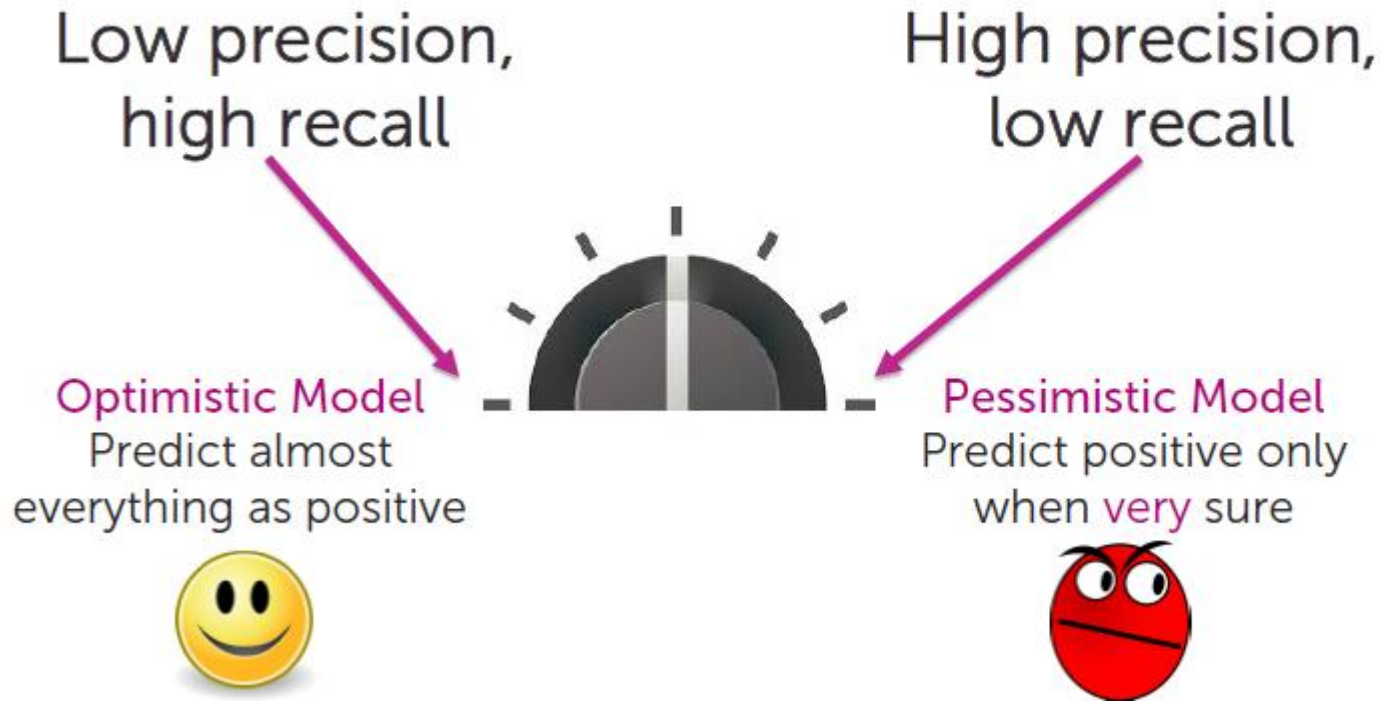
High precision, low recall



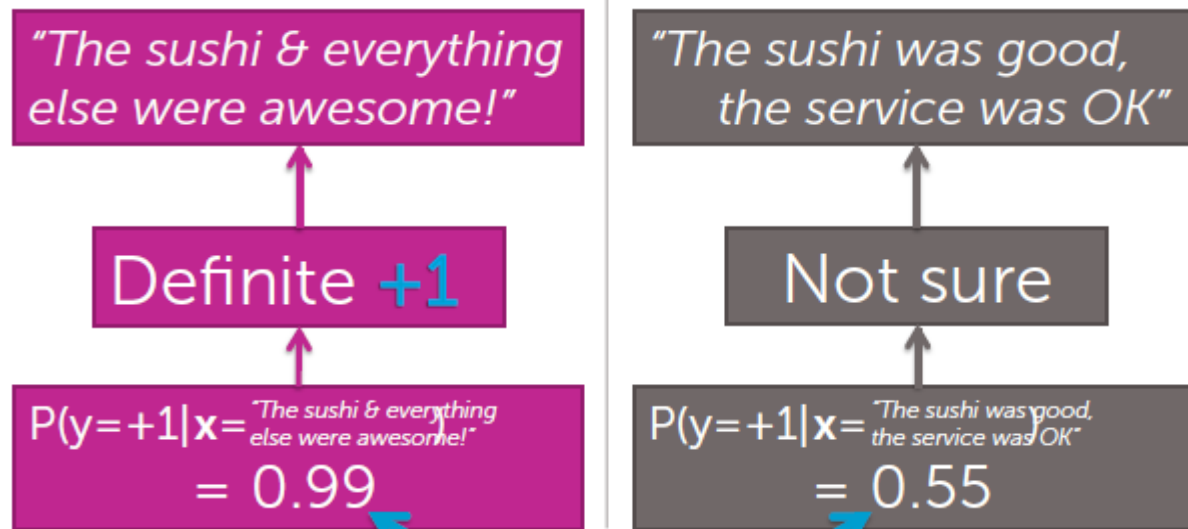
# Balancing precision & recall



# Can we tradeoff precision & recall?

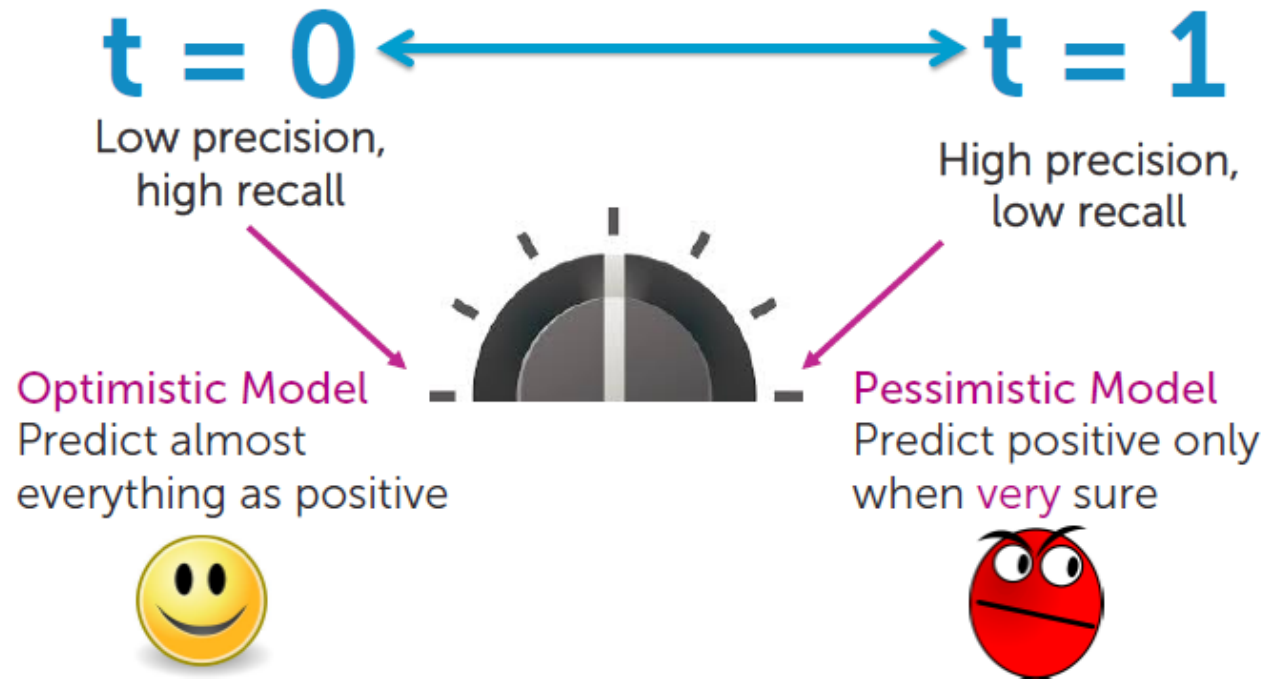


## How confident is your prediction?



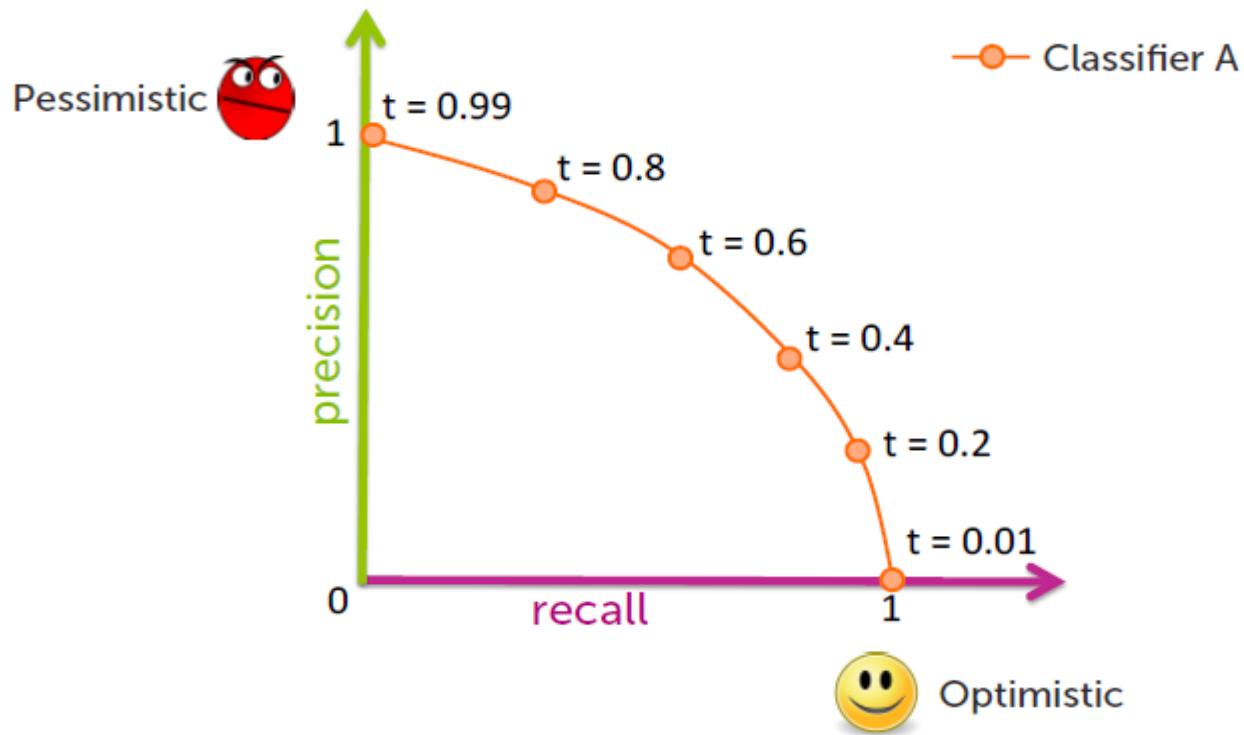
Can be used to tradeoff precision and recall

## Tradeoff precision & recall with threshold

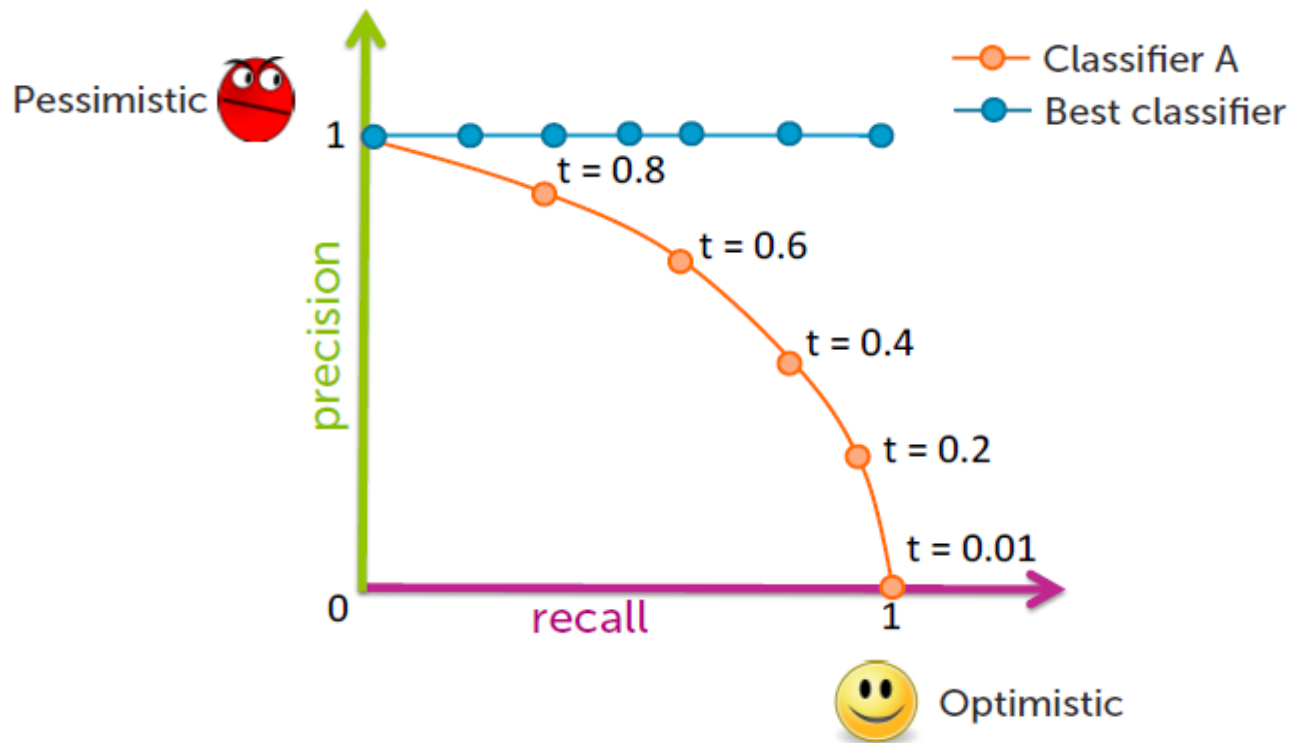




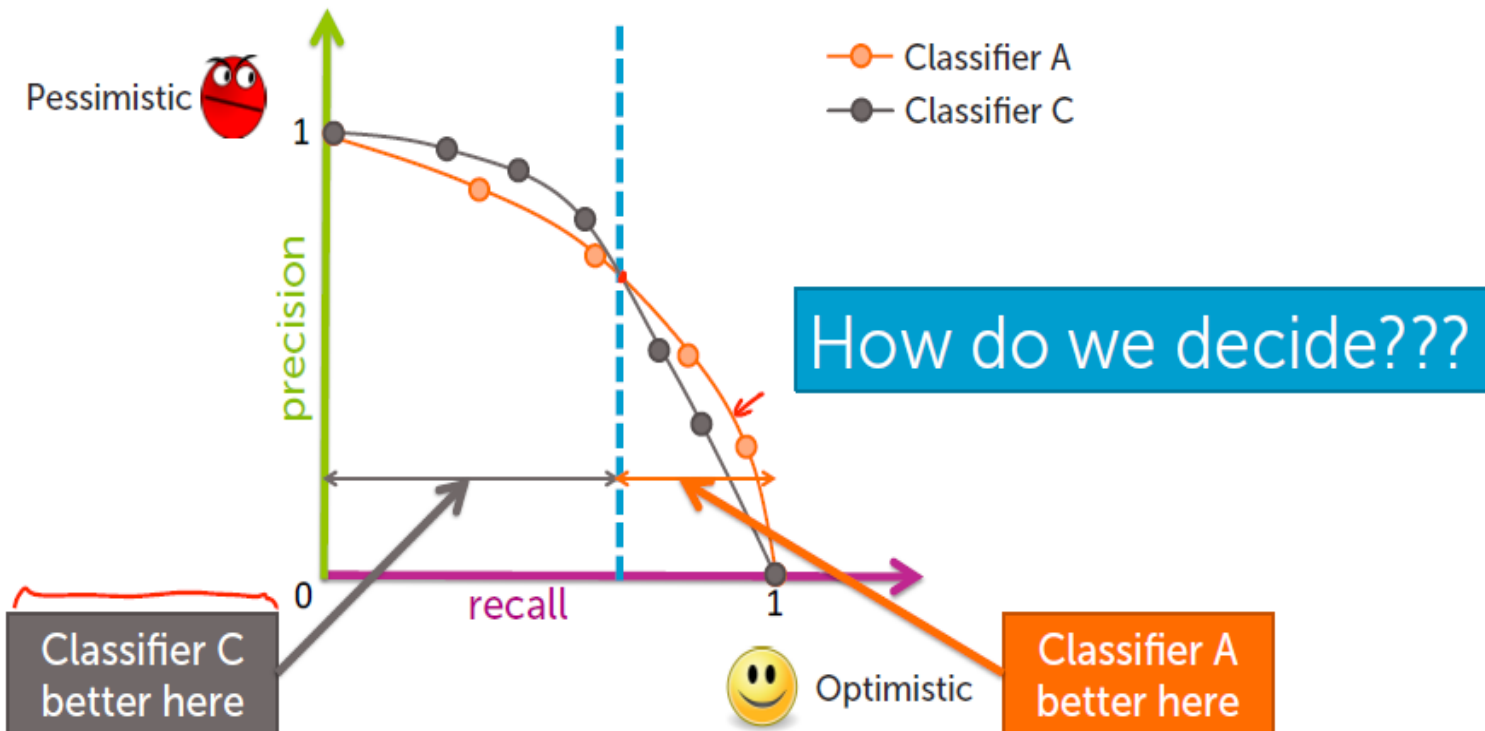
# The precision-recall curve



## What does the perfect algorithm look like?



# Which classifier is better? A or C?



# Compare algorithms

- Often, reduce precision-recall to single number to compare algorithms
  - F1 measure, area-under-the-curve (AUC),...

## Precision at k

Showing  
k=5 sentences  
on website



Sentences model  
most sure are positive

Easily best sushi in Seattle.	
My wife tried their ramen and it was pretty forgettable.	
The sushi was amazing, and the rice is just outstanding.	
All the sushi was delicious.	
The service was perfect.	



precision at k = 0.8

# What you can do now

29

- Classification accuracy/error are not always right metrics
- **Precision** captures fraction of positive predictions that are correct
- **Recall** captures fraction of positive data correctly identified by the model
- Trade-off **precision** & **recall** by setting probability thresholds
- Plot **precision-recall** curves.
- Compare models by computing precision at **k**