

Unranked retrieval

- Precision:
 - Fraction of retrieved documents that are relevant
- Recall:
 - Fraction of relevant documents that are retrieved



Unranked retrieval

- Precision:
 - Fraction of retrieved documents that are relevant
- Recall:
 - Fraction of relevant documents that are retrieved

	<i>Relevant</i>	<i>Not Relevant</i>
<i>Retrieved</i>	<i>TP</i>	<i>FP</i>
<i>Not Retrieved</i>	<i>FN</i>	<i>TN</i>



Unranked retrieval

- Precision:
 - Fraction of retrieved documents that are relevant
- Recall:
 - Fraction of relevant documents that are retrieved

	<i>Relevant</i>	<i>Not Relevant</i>
<i>Retrieved</i>	<i>TP</i>	<i>FP</i>
<i>Not Retrieved</i>	<i>FN</i>	<i>TN</i>

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$



Unranked retrieval - Accuracy

- The difficulty with measuring “accuracy”
- In one sense accuracy is how many judgments you make correctly

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

	<i>Relevant</i>	<i>Not Relevant</i>
<i>Retrieved</i>	<i>TP</i>	<i>FP</i>
<i>Not Retrieved</i>	<i>FN</i>	<i>TN</i>

- Why is this not a very useful measure?



Exercise

- Documents A - F, Query q

<i>Document</i>	<i>Relevant</i> (q)	<i>Not Relevant</i> (q)
<i>A</i>	✓	
<i>B</i>		✓
<i>C</i>		✓
<i>D</i>	✓	
<i>E</i>		✓
<i>F</i>	✓	

- If my system returns A,C,D,E to query q
 - How many TP, TN, FP, FN do I have?

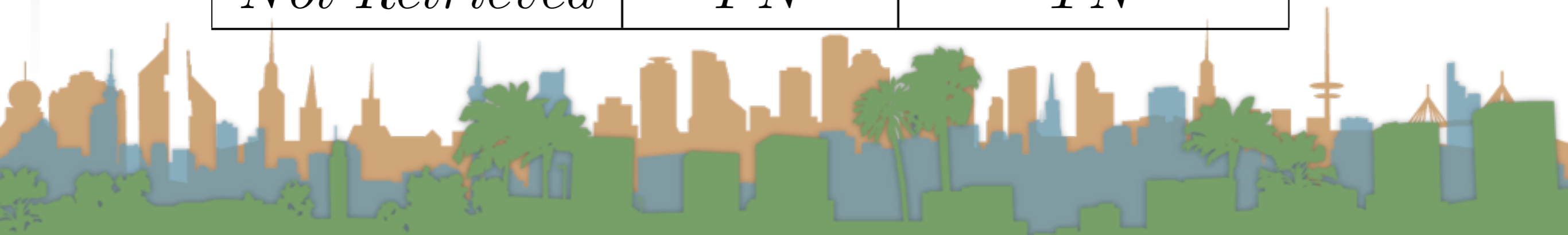


Exercise

Retrieved : A C D E

<i>Document</i>	<i>Relevant(q)</i>	<i>Not Relevant(q)</i>
<i>A</i>	✓	
<i>B</i>		✓
<i>C</i>		✓
<i>D</i>	✓	
<i>E</i>		✓
<i>F</i>	✓	

	<i>Relevant</i>	<i>Not Relevant</i>
<i>Retrieved</i>	<i>TP</i>	<i>FP</i>
<i>Not Retrieved</i>	<i>FN</i>	<i>TN</i>



Exercise

Retrieved : A C D E

<i>Document</i>	<i>Relevant(q)</i>	<i>Not Relevant(q)</i>
<i>A</i>	✓	
<i>B</i>		✓
<i>C</i>		✓
<i>D</i>	✓	
<i>E</i>		✓
<i>F</i>	✓	

	<i>Relevant</i>	<i>Not Relevant</i>
<i>Retrieved</i>	<i>TP</i>	<i>FP</i>
<i>Not Retrieved</i>	<i>FN</i>	<i>TN</i>



Exercise

- What is our precision?

$$\textit{Precision} = \frac{TP}{TP + FP}$$

<i>TP</i>	2
<i>FP</i>	2
<i>FN</i>	1
<i>TN</i>	1

- What is our recall?

$$\textit{Recall} = \frac{TP}{TP + FN}$$

- What is our accuracy?

$$\textit{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}$$



Exercise

- If my system returns A,C,D,E to query q....

<i>Document</i>	<i>Relevant(q)</i>	<i>Not Relevant(q)</i>
<i>A</i>	✓	
<i>B</i>		✓
<i>C</i>		✓
<i>D</i>	✓	
<i>E</i>		✓
<i>F</i>	✓	

<i>Precision</i>	$\frac{1}{2}$
<i>Recall</i>	$\frac{2}{3}$
<i>Accuracy</i>	$\frac{1}{2}$

- What do I want Precision to be?

	<i>Relevant</i>	<i>Not Relevant</i>
<i>Retrieved</i>	<i>TP</i>	<i>FP</i>
<i>Not Retrieved</i>	<i>FN</i>	<i>TN</i>

$$Precision = \frac{TP}{TP + FP}$$



Exercise

- If my system returns A,C,D,E to query q....

<i>Document</i>	<i>Relevant(q)</i>	<i>Not Relevant(q)</i>
<i>A</i>	✓	
<i>B</i>		✓
<i>C</i>		✓
<i>D</i>	✓	
<i>E</i>		✓
<i>F</i>	✓	

<i>Precision</i>	$\frac{1}{2}$
<i>Recall</i>	$\frac{2}{3}$
<i>Accuracy</i>	$\frac{1}{2}$

- What do I want Recall to be?

	<i>Relevant</i>	<i>Not Relevant</i>
<i>Retrieved</i>	<i>TP</i>	<i>FP</i>
<i>Not Retrieved</i>	<i>FN</i>	<i>TN</i>

$$Recall = \frac{TP}{TP + FN}$$



Exercise

- If my system returns A,C,D,E to query q....

<i>Document</i>	<i>Relevant(q)</i>	<i>Not Relevant(q)</i>
<i>A</i>	✓	
<i>B</i>		✓
<i>C</i>		✓
<i>D</i>	✓	
<i>E</i>		✓
<i>F</i>	✓	

Precision

$\frac{1}{2}$

Recall

$\frac{2}{3}$

Accuracy

$\frac{1}{2}$

- What do I want Accuracy to be?

	<i>Relevant</i>	<i>Not Relevant</i>
<i>Retrieved</i>	<i>TP</i>	<i>FP</i>
<i>Not Retrieved</i>	<i>FN</i>	<i>TN</i>

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$



Unranked retrieval - Accuracy

- Welcome to my search engine
- I guarantee a 99.9999% accuracy.
- Bring on the venture capital

Beta

PITTERPATTERSONFINDER

Search for:



Unranked retrieval - Accuracy

- Welcome to my search engine
- I guarantee a 99.9999% accuracy.
- Bring on the venture capital

Beta

PITTERPATTERSONFINDER

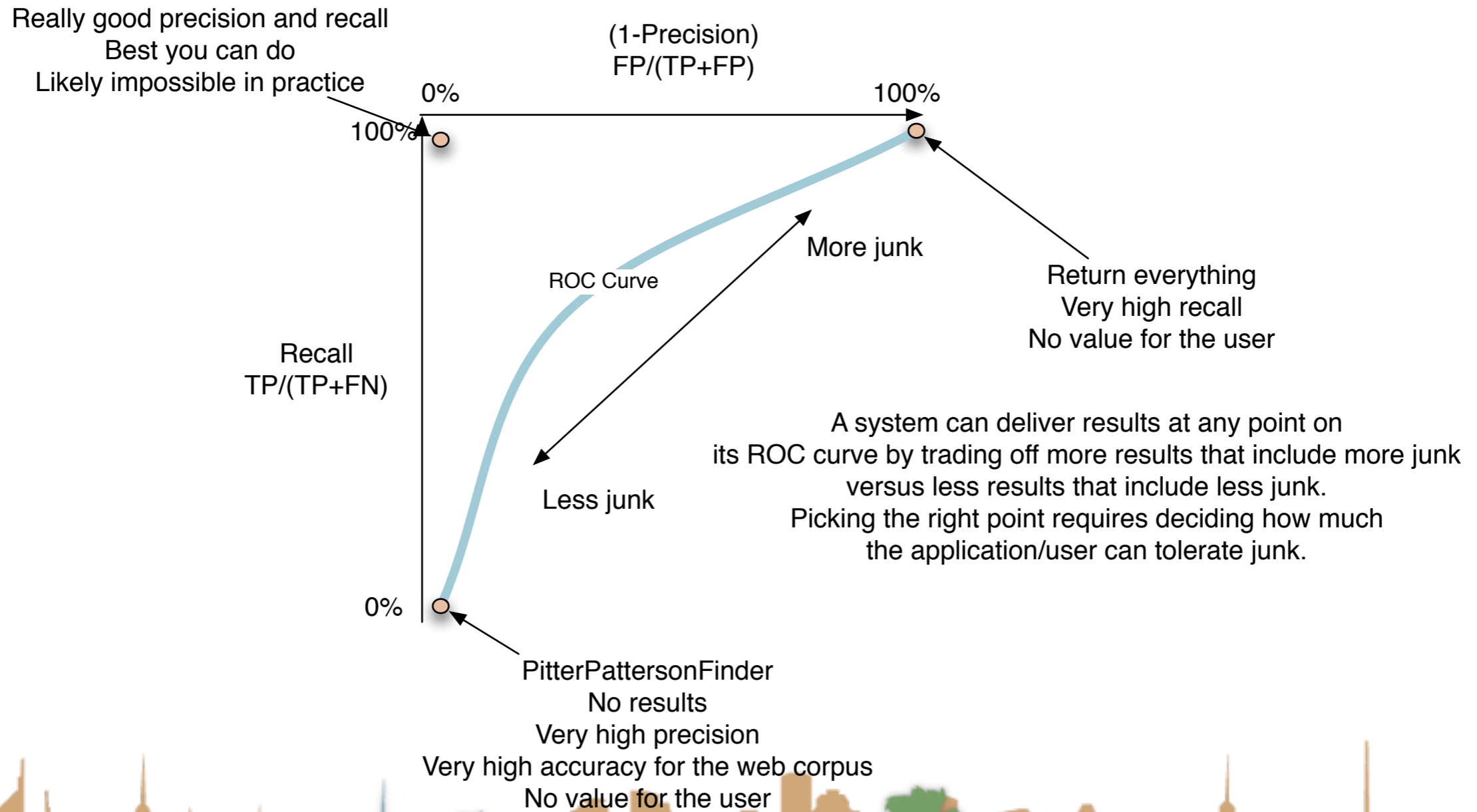
Search for:

- Most people want to find something and can tolerate some junk

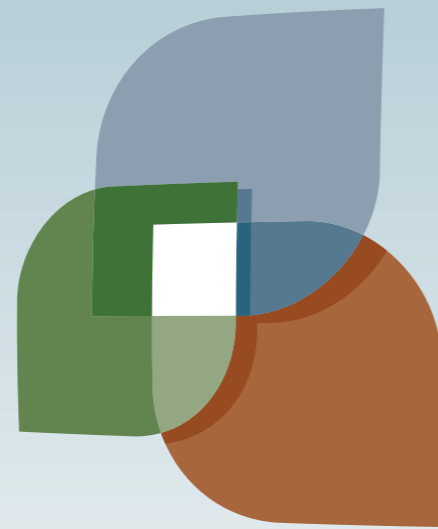


Unranked retrieval - ROC curve

Receiver Operating Characteristic (ROC) curve



next...



L U C I

