Information Retrieval (Relevance Feedback & Query Expansion)

Fabio Aiolli

http://www.math.unipd.it/~aiolli

Dipartimento di Matematica Pura ed Applicata Università di Padova

Anno Accademico 2008/2009

Dip. di Matematica Pura ed Applicata

Pura ed Applicata

F. Aiolli - Information Retrieval 2008/2009

1

Relevance feedback and guery expansion Goal: To refine the answer set by involving the user in the retrieval process (feedback/interaction) Local Methods (adjust the user queries) Relevance feedback Pseudo (or Blind) Relevance Feedback (Global) indirect Relevance Feedback Global Methods (independent of the gueries and results) Query expansion/reformulation with a thesaurus (WordNet) Query expansion via automatic thesaurus generation Other techniques (spelling correction,...) 2 Dip. di Matematica F. Aiolli - Information Retrieval

2008/2009

Rele	evance	теедраск	
□ Bo 1. 2. 3.	sic Procedu The user is The system The user m	re of RF sues a simple query n returns an initial set of retrieval resu arks some of these documents as	ılts
4. 5.	The system information The system	n computes a better representation of n need based on this feedback n displays a revised set of results	the
Re	peat the pr	ocedure one or more times.	
□ Th inf	is process formation ne	helps the user to focalize its own eed as well.	
Dip. di Mat	tematica	F. Aiolli - Information Retrieval	3
Rele	2VANCE nage sear	Feedback: Example	
Rele	EVANCE nage sear 'tp://nayal	Feedback: Example The Engine na.ece.ucsb.edu/imsearch/imsea	rch.htm _□□×
Rele	EVANCE nage sear tp://nayal	Feedback: Example och engine na.ece.ucsb.edu/imsearch/imsear pe Bookmarks Iools Window Help	rch.htm
Rele	EVANCE nage sear tp://nayal New Page 1 - Netsca Ele Edit View Go	Feedback: Example The engine na.ece.ucsb.edu/imsearch/imsear Pe Bookmarks Iools Window Help Cook Mindow Help	rch.htm -□×
Rele	EVANCE nage Sear tp://nayal New Page 1 - Netsca Elle Edit View Go	Feedback: Example rch engine na.ece.ucsb.edu/imsearch/imsear pe Bookmarks Iools Window Help Contempo Inter://nayana.ece.ucsb.edu/i	rch.htm -□× €
Rele	EVANCE nage sear tp://nayal New Page 1 - Netsca Elle Edit View Go	Feedback: Example rch engine na.ece.ucsb.edu/imsearch/imsear Pe Bookmarks Iools Window Help Intp://nayana.ece.ucsb.edu/i	rch.htm -□× €
Rele	EVANCE nage Sear tp://nayal New Page 1 - Netsca Elle Edit View Go Elle Edit View Go Home & Browsir Shopping relat	Feedback: Example rch engine na.ece.ucsb.edu/imsearch/imsear pe Bookmarks Iools Window Help Point Inter://nayana.ece.ucsb.edu/i	rch.htm -□× €
Rele	EVANCE nage sear tp://nayai New Page 1 - Netsca Elle Edit View Go Elle Edit View Go	Feedback: Example The engine na.ece.ucsb.edu/imsearch/imsear pe Bookmarks Iools Window Help Point Intp://nayana.ece.ucsb.edu/i Intp://nayana.ece.ucsb.	rch.htm × e 1000)



Relevance Feedback step

			Diowse		
	Bitters	J.J.		EIKING 2000 BIKE VEXE	
44473, 16458)	(144457, 252140)	(144456, 262857)	(144456, 262863)	(144457, 252134)	(144483, 265154)
.0	0.0	0.0	0.0	0.0	0.0
.0	0.0	0.0	0.0	0.0	0.0
độ			J.		3 A
	(144483, 265153)	(144518, 257752)	(144538, 525937)	(144456, 249611)	(144456, 250064)
144483, 264644)		0.0	0.0	0.0	0.0
144483,264644) .0	0.0	0.0	0.0		

Results after Relevance Feedback



Rocchio Algorithm

- □ The Rocchio algorithm incorporates relevance feedback information into the vector space model.
- □ Want to maximize $sim(Q, C_r) sim(Q, C_{nr})$
- The optimal query vector for separating relevant and non-relevant documents (with cosine sim.):



- **Q**_{opt} = optimal query; C_r = set of rel. doc vectors; N = collection size
- Unrealistic: we don't know relevant documents.

Dip. di Matematica Pura ed Applicata













Dip. di Matematica Pura ed Applicata F. Aiolli - Information Retrieval 2008/2009











Automatic Thesaurus Generation Example

word	ten nearest neighbors
absolutely	absurd whatsoever totally exactly nothing
bottomed	dip copper drops topped slide trimmed slig
captivating	shimmer stunningly superbly plucky witty
doghouse	dog porch crawling beside downstairs gazed
Makeup	repellent lotion glossy sunscreen Skin gel pe
mediating	reconciliation negotiate cease conciliation p
keeping	hoping bring wiping could some would oth ϵ
lithographs	drawings Picasso Dali sculptures Gauguin 1
pathogens	toxins bacteria organisms bacterial parasit ϵ
senses	grasp psyche truly clumsy naive innate awl

Dip. di Matematica Pura ed Applicata F. Aiolli - Information Retrieval 2008/2009

