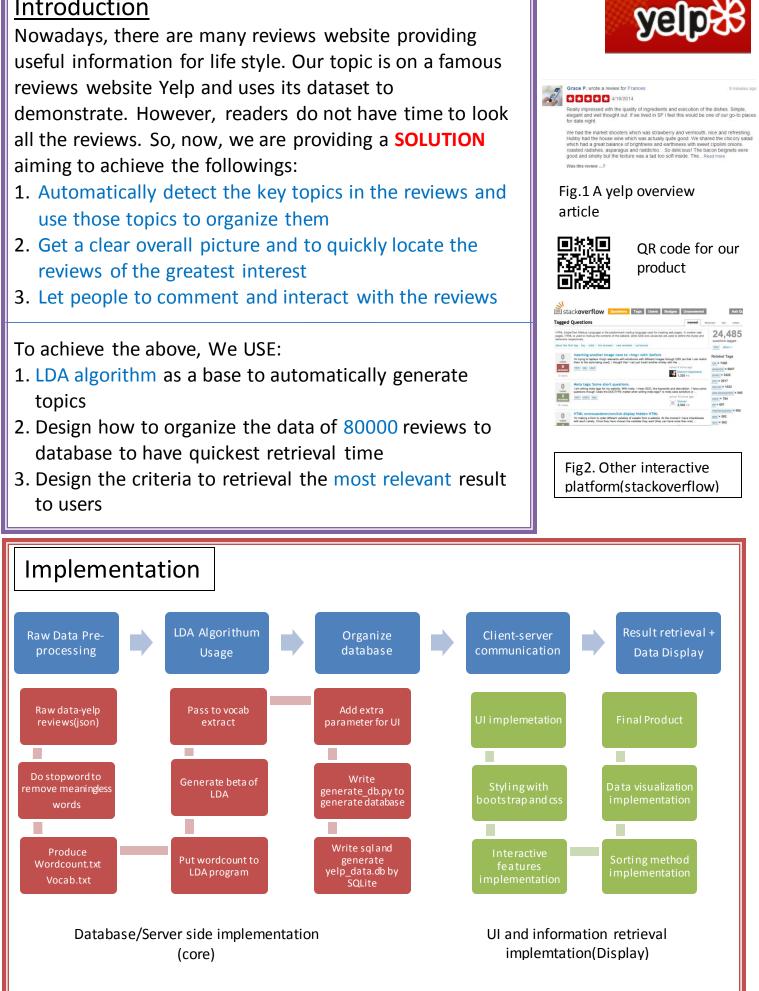


interesting Interactive Online Platform For Discussing Yelp **Reviews Grouped By Topics Using Topic-Modeling** Group Member: Chan Ki Yung, Lam Yuen Ting, Li Chun Ho Supervisor: Prof. Nevin L. Zhang





Introduction



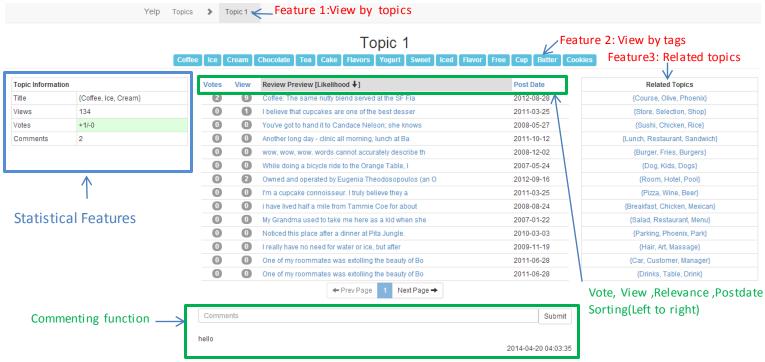


Fig3. Overview of UI in Topic 1

Features:

Our goal:

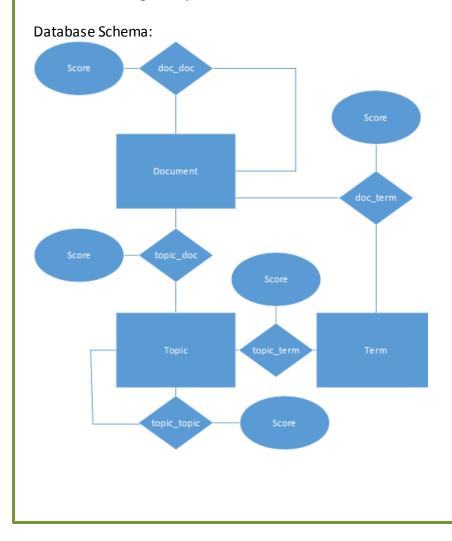
- 1. Grab the first sight idea of the topic
- 2. Get the statistic of overview of the dataset
- 3. To let views to interact with reviews

Result retrieval Features: (in red)

-"Topic" Features (Deep blue box) 15 large categories of topics to let viewers grab big picture -"Tag" Feature -"Relative topic" Feature Interactive Features: (in green) 1) Vote: +/- vote 2) Number of views 3) Comments (comment on reviews) 4) Sort by postdate/relevance Statistics features: 1) Overall views 2) Number of article in topic/tag 3) Total +/- vote 4) Number of comments (hot topic) 5) Postdate extraction

Database Design:

One of the notable part of this project is on how to organize the database to give a quick retrieval.



Evaluation Testing Results:

Client side testina:

Chefit side testing.		-	1	
ltem	Passing Criteria	Result	Conclusion	
Data Retrieval -Correctness	100%	100%	Pass	
Capacity of database	>40000 documents	80000	Pass	
Server Client interaction -hyperlinks	100% to desired pages	100%	Pass	
Server Side Testing:				
ltem	Passing Criteria	Result	Conclusion	
Pre-processing program	100%match	100%	Pass	
Data loss betw een conversion	0%	0%	Pass	
Information consistency	100%	100%	Pass	

Speed Test (topic page generation):

From Dulles,VA,US under IE8 by webpagetest.org

domContentLoaded lossEvent 4288 - 4278 (0.0115) 4208 - 4283 (0.0035) Uterfall View Content Download Sursegonse Content Download DNS Loolup Initial Contention Time to First Byte Content Download Sursegonse Content Download Marginguese unit.Hs;1722-rbsic.download 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.0 2.0 2.1 2.4 2.6 2.0 3.0 3.2 3.4 3.6 3.8 4.6 1. Montguese unit.Hs;1722 - nain.ces 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.0 2.0 2.1 2.4 2.6 2.0 3.0 3.2 3.4 3.6 3.8 4.6 1. Montguese unit.Hs;1722 - nain.ces 0.0 0.0 1.0 1.2 1.4 1.6 1.0 2.0 2.1 4.4 2.6 2.0 3.0 3.2 3.4 3.6 3.8 4.6 0. retrain.botters 0.0 0.0 1.0 1.2 1.4 1.6 1.0 2.0 2.1 4.4 2.6 2.0 3.0 3.2 3.4 3.6 3.8 4.6 0. retrain.botters 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.1 4.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.6 0. retrain.botters 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.1 4.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.6 0. retrain.botters 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.1 4.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.6 0. retrain.botters 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.1 4.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.6 0. retrain.botters 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.0 2.4 4.2 6.2 0.3 0.3 2.3 4 3.6 3.8 4.6 0. retrain.botters 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0	Time	First Byte	Start Render	Visually Complete	Speed Index	DOM Elements	Result (error code)	Time	Requests	Bytes In	Time	Requests	Bytes In
1 1288 + 278 (0.013) 1289 + 233 (0.003) Waterfall View Image: State of the st	.255s	1.756s	2.257s	2.400s	2301	184	99999	2.255s	7	184 KB	2.536s	8	184 KB
1 1288 + 278 (0.013) 1289 + 233 (0.003) Waterfall View Image: State of the st					domCont	entl oaded	loadEvent						
Waterfail View DNS Lookup Initial Connection Time to First Brie Content Download Sxr response Rest response Listar Render DDM Content Loaded On Load Document Complete Iter/Audrigu.csc.uet.18:7222-response Rest response Rest response<								•					
DNS Lookup Initial Connection Time to First Byte Content Download Sorr response Reservember Istan Render DOM Content Loaded On Load Document Complete 1: Multipactes unit.NE/1227 typic.ph 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.0 2.0 2.1 2.4 2.6 2.0 3.0 3.2 3.4 3.6 3.8 4.6 1.0 1.2 0.0 1.2 1.9 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0					4.2003 - 4.2	.100 (0.0110)	1.2003 - 4.2033 (0.0032	"					
DNS Lookup Initial Connection Time to First Byte Content Download Sorr response Reservember Istan Render DOM Content Loaded On Load Document Complete 1: Multipactes unit.NE/1227 typic.ph 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.0 2.0 2.1 2.4 2.6 2.0 3.0 3.2 3.4 3.6 3.8 4.6 1.0 1.2 0.0 1.2 1.9 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0						Motorfo	II Man						
IstantRender DOM Content Loaded On Load Document Complete ttgs://www.gp.cces.ust.tki?122?-tspic.g 0,2 0,4 0,4 0,4 1.4 1.4 1.0 2.0 2.2 2.4 2.6 2.0 3.0 3.2 3.4 3.6 3.8 4.6 1. www.gp.cces.ust.tki?122?-tspic.g 0.2 0.4 0.4 0.4 1.4 1.4 1.0 2.0 2.2 2.4 2.6 2.0 3.0 3.2 3.4 3.6 3.8 4.6 2. www.gp.cces.ust.tki?122 restinx.boststraboststrag.in.css 9.0						valena	ii view						
ttp://wwfg.csc.ust.Wr/122/Topics 1. uwfg.csc.ust.Wr/122/Topics 2. uwfg.csc.ust.Wr/122/Topics 2. uwfg.csc.ust.Wr/122/Topics 2. uwfg.csc.ust.Wr/122/Topics 2. uwfg.csc.ust.Wr/122/Topics 2. uwfg.csc.ust.Wr/122/Topics 3. ustan.boottraboottra 5. ustan.boottraboottra 5. ustan.boottra 5. ustan.boottra			DNS Loo	kup 📕 Initial Con	nection 📃 1	ime to First Byte	Content Downloa	ad 3xx I	esponse	4xx+ resp	onse		
ttp://wwfg.csc.ust.Wr/122/Topics 1. uwfg.csc.ust.Wr/122/Topics 2. uwfg.csc.ust.Wr/122/Topics 2. uwfg.csc.ust.Wr/122/Topics 2. uwfg.csc.ust.Wr/122/Topics 2. uwfg.csc.ust.Wr/122/Topics 2. uwfg.csc.ust.Wr/122/Topics 3. ustan.boottraboottra 5. ustan.boottraboottra 5. ustan.boottra 5. ustan.boottra													
1. workpace.ust.NR:722 - topic.org 2. workpace.ust.NR:722 - topic.org 3. workpace.ust.NR:722 - topic.org 4. workpace.ust.NR:722 - topic.org 4. workpace.ust.NR:722 - topic.org 4. workpace.ust.NR:722 - topic.org 5. workpace.ust.NR:722 - topic.org 6. workpace.ust.NR:722 - topic.org 7. workpace.ust.				Start Rend	er DOM (Content Loaded	On Load Docu	ment Co	mplete				
1. workpace.ust.NR:722 - topic.org 2. workpace.ust.NR:722 - topic.org 3. workpace.ust.NR:722 - topic.org 4. workpace.ust.NR:722 - topic.org 4. workpace.ust.NR:722 - topic.org 4. workpace.ust.NR:722 - topic.org 5. workpace.ust.NR:722 - topic.org 6. workpace.ust.NR:722 - topic.org 7. workpace.ust.													
2. workgo.ccs.ut.k1212 - nainces a. refami.bottarsbottars.nin.s. 4. refami.bottarsbottars.nin.s. 5. refami.bottarsbottars.nin.s. 6. refami.bottarsbott					.4 0.6 0.8	1.0 1.2 1.				3.0 3.2	3.4	3.6 3.8	1.0 4.2
3. referminototrabootrane.in.cos 4. referminototrabootrane.in.cos 4. referminototrabootrane.in.cos 4. referminototrainge-regular outrabootrane.inge- 5. cose.jugery.comerg-1.01.cl.nin.js 6. referminototrainge-regular outrabootrane.inge- 7. Marija.cde unt.ing-7122 - Fauton.ico 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.0 2.0 2.2 2.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.0 CFU Utilization							50						
4. referant southers toottersmains 5. referant southers trips-regular work 6. referant southers trips-regular work 6. referant southerstrips-regular work 6. referant southerstrips-referant southers													
6. retrini.hotetraling=-regular.uoF 7. wwfgc.cse.uut.Hs7122 - ForCon.ico 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.0 3.0 3.2 3.4 3.6 3.8 4.4 CPU Utilization													
7. Mor/paces units/322 - Feuton.ico 7. Mor/paces units/322 - Feuton.ico 7. Mor/paces units/322 - Feuton.ico 0.2 0.4 0.5 0.0 1.0 1.2 1.4 1.6 1.0 2.0 2.2 2.4 2.6 2.0 3.0 3.2 3.4 3.6 3.8 4.6 CFU Utilization	code.jq	uery.com.	uery-1.10.1	.min.js			218	ns					
0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.6	. netona.	bootstra.	lings-regul	ar.woff				60 MS					
CFU Utilization	NUMFYD.	cse.ust.hi	k:7122 - favi	con.ico					257 MS (4	04)			
				0.2 0	.4 0.6 0.8	1.0 1.2 1.	4 1.6 1.8 2.0 2.	2 2.4	2.6 2.8	3.0 3.2	3.4	3.6 3.8	4.0 4.2
							1						
	- CPU I	Utilizati	on					\mathbf{N}					
						~	~,		\sim	\sim	_	_	\sim
	Daniel						\wedge						
Bandardonin (0 - 5,000 kops)	BandwidthIn (0 - 5,000 Kbps)		\$)										
							· · · · ·				_		

What's Next?



Search engine technique?



Social network?



forum?

Result and Future Work

Future work:

- 1. Continue with the project and build a community to rate on reviews
- 2. Use LDA to analyze the correlation between positive/negative wordings and the stars/vote of an article

Result:

The result of this project is satisfactory. The success of this project gives an alternative way for information retrieval other than search engine. The application of LDA to information retrieval site can greatly reduce the documents which are irrelevant to the topics and easier for user to search the article desired. This may be the solution to information retrieval in the era of information explosion.