

Evaluating_rank_results_v2

June 15, 2018

1 Evaluation des DESM-Ranking

```
In [2]: import pandas as pd
import numpy as np
```

Zuerst laden wir den Korpus "News Vertical Search Dataset" als pandas Dataframe in den Speicher (in Variable "corpus"). Wir entfernen unnötige Spalten.

```
In [3]: corpus = pd.read_csv("../data/eval_corpus.csv", sep=";")
del corpus['Unnamed: 0']
```

Schauen wir uns die existierenden Suchanfragen an und wählen zufällig Anfragen die wir testen wollen.

```
In [4]: corpus.shape
```

```
Out[4]: (830, 5)
```

```
In [17]: set(corpus["query"])
```

```
Out[17]: {'a_hoax',
'a_video',
'aaron_is_dead',
'aaron_swartz',
'amber_alert',
'american_airlines',
'angelina_jolie',
'armed_guards',
'big_board',
'blue_jackets',
'cam_newton',
'cartridges_didn't',
'ces_2013',
'chase_jarvis',
'chip_kelly',
'cirque_du_soleil',
'cnn_chile',
'colonial_marines',
```

'david_miscavige',
'debt_ceiling',
'devil_may_cry',
'duchess_of_cambridge',
'el_riesgo_de',
'fc_barcelona',
'former_bodyguard',
'fue_el',
'golden_globes',
'gun_control',
'gun_violence',
'i_know',
'ice_cream',
'indiana_jones',
'internet_genius',
'iphone_5s',
'issued_for',
'jack_lew's_signature",
'jodie_foster',
'justin_timberlake',
'katherine_webb',
'kris_humphries',
'la_mega',
'lance_armstrong',
'lena_dunham',
'lew's_signature_is",
'met_office',
'mike_mccoy',
'miss_america',
'new_maps',
'nfl_kicks',
'para_la',
'penny_arcade',
'return_to_music',
'roger_ebert',
'ryan_seacrest',
'school_shooting',
'serious_jibber-jabber',
'set_for',
'sky_sports',
'spring_breakers',
'super_bowl',
'team_sky',
'the_national_anthem',
'the_nfl',
'the_nhl',
'the_tech',
'tom_brady',

```
'training_camp',
'un_tren',
'van_persie',
'video_of_the_day',
'violent_video_games',
"what's_the_point",
'white_house',
'whole_foods_ceo'}
```

1.1 Anfrage 1

Die erste Suchanfrage, die wir testen ist "what's_the_point". Im folgenden wird pro Anfrage die gleiche Struktur verwendet: Als erstes schauen wir uns die (laut Test-Korpus) relevanten Dokumente in einer Uebersicht an. Darauf folgen die Ergebnisse (Ranking-Scores und entsprechende Indices sowie headlines) des DESM-Rankings auf dem gesamten Korpus. Zum Schluss folgt jeweils eine Beschreibung der Beobachtungen von den Ergebnissen.

```
In [18]: # query: whats the point
corpus[corpus["query"] == "what's_the_point"]
```

```
Out[18]:
```

	source	query	score	\
550	reddit	what's_the_point	5	
551	news	what's_the_point	1	
552	news	what's_the_point	1	
553	news	what's_the_point	1	
554	reddit	what's_the_point	14	
555	news	what's_the_point	1	
556	reddit	what's_the_point	54	
557	news	what's_the_point	1	
558	reddit	what's_the_point	22	
559	news	what's_the_point	1	

	headline	\
550	If you're an atheist, what's the point in livi...	
551	FoxBusiness.com - Time to Pay...	
552	The Oxford Eagle - Jimmy Reed	
553	Berkshire Eagle - Man charged...	
554	At what point does it stop being coffee?	
555	LA Times - December home pric...	
556	At what point do I disclose this somewhat cont...	
557	How To - Hit a Scissor Kick Overhead - ...	
558	Follow Up To: At what point do you just Give U...	
559	Another Day, Another Goldman Sachs Exec, Anoth...	

	summary
550	Before the flaming starts, read what I have to...
551	You'll want to get to the ...
552	Local columnist Jimmy Reed writes about a youn...

553 A Guatemalan man living in Pittsfield is being...
 554 A lot of people say they like coffee, but what...
 555 The median is the point</stron...
 556 I know this belongs in /r/askmen but this sub ...
 557 When you're at net and a lob goes over you...
 558 Four days ago I posted [At what point do you j...
 559 (On this count, the point</str...

Ranking Resultate

Score: -0.014798
 Score: -0.015701
 Score: -0.016491
 Score: -0.016642
 Score: -0.017701
 Score: -0.018240
 Score: -0.018658
 Score: -0.018713
 Score: -0.018824
 Score: -0.018914
 Score: -0.019741
 Score: -0.019768
 Score: -0.019769
 Score: -0.019800
 Score: -0.019886
 Score: -0.020071
 Score: -0.020086
 Score: -0.020126
 Score: -0.020245
 Score: -0.020331

557 How To - Hit a Scissor Kick Overhead - tennis.com
 465 Serious question: When can people "pirate" airline tickets?"
 554 At what point does it stop being coffee?
 558 Follow Up To: At what point do you just Give Up"? "
 24 Messi renews Barcelona contract - Taipei Times
 101 What's the weather like where you are?
 819 Is ironman mode a new concept in gaming that I've missed until now?
 129 Philly.com - The Numbers Behind Billick - BoopStats
 339 Pats hope to continue success without Gronkowski - NewsOK.com
 290 Super Bowl Predictions: Must include justification.
 523 ContraCostaTimes.com - Colin Kaepernick's college coach to his pupil - 'Just awesome' - ContraCostaTimes.com
 464 2012 MacBook Pro internet is spotty.
 194 Known hoaxes?
 499 Can someone please explain to me what the point of having a PAT" is and not just a straight 7 point touchdown?"
 608 A Fresh Approach - Mailtribune.com
 497 NFL To Kick Off Exciting Playoff Weekend - WKAR
 109 CNN.com - Ice, fog spread across United Kingdom - Photos - CNN.com

708 Are Video Game Shows Dying Out

559 Another Day, Another Goldman Sachs Exec, Another Gwathmey Penthouse Sale - Observer

618 Austin's source for local news - Joystick takes an early look at Deadpool game - The Austin Daily Herald

Get position of relevant Documents in Ranking:

550 21

551 34

552 97

553 388

554 2

555 91

556 347

557 0

558 3

559 18

Median of Ranking-Result: 27.5

In den Top 20 Ergebnissen sind 4 von 10 nach dem Testkorpus relevanten Dokumente vertreten. Anscheinend besteht ein groeSSerer Teil der Ergebnisdokumente aus Fragestellungen, was die Query in ihrer Form als Frage widerspiegelt (vielleicht auch durch das Wort "what"). Im folgenden Beispielhaft der Platz 2 im Ranking, der nicht in den relevanten Dokumenten ist.

```
In [19]: corpus.ix[465]["summary"]
```

```
Out[19]: "It's all connected to the internet. Some genius out there can find a bug to exploit. I
```

1.2 Anfrage 2

```
In [20]: # Query: aaron is dead
         corpus[corpus["query"] == "aaron_is_dead"]
```

```
Out[20]:
```

	source	query	score	\
590	reddit	aaron_is_dead	1	
591	reddit	aaron_is_dead	7	
592	reddit	aaron_is_dead	10	
593	reddit	aaron_is_dead	5	
594	reddit	aaron_is_dead	96	
595	reddit	aaron_is_dead	12	
596	reddit	aaron_is_dead	9	
597	reddit	aaron_is_dead	5	
598	reddit	aaron_is_dead	10	
599	reddit	aaron_is_dead	9	

	headline	\
590	Who should I pick up from the waiver wire: Cra...	
591	Laying a Trap Part 2	
592	Favorite Fight Friday: Aaron Pryor vs Alexis A...	
593	[KEEPER ADVICE] - At what point do you throw i...	
594	Aaron Shwartz is dead - Reddit, we did it!	
595	So is Bruce Banner Back? (Incredible Hulk #7.1)	

596 [Story 1] The Greater Thelema Society - A Disp...
 597 Aaron's Android Advice reviews: Mage Defense
 598 Name 3 people dead or alive that you would lov...
 599 [Contains Huge Comic Spoilers]: Fan speculatio...

summary

590 I plan on dropping Victor Cruz for one of the...
 591 2nd Squadron surged from the Thermopylae, secu...
 592 With all the resurgence of the Pacweather talk...
 593 Hey all. In my PPR league, I'm now 3-6 and dea...
 594 .
 595 So 7 issues in and Bruce Banner is already bac...
 596 Its about time you showed up.\n\nDont act...
 597 Aarons Android Advice\n\nTITLE:\tMage Defense...
 598 My room mate and I got vaped tonight and decid...
 599 So I was having a discussion about TWD with th...

Ranking Resultate

Score: -0.020255
 Score: -0.034791
 Score: -0.034818
 Score: -0.036768
 Score: -0.037668
 Score: -0.038481
 Score: -0.039062
 Score: -0.039318
 Score: -0.039349
 Score: -0.039872
 Score: -0.041260
 Score: -0.041516
 Score: -0.041943
 Score: -0.042028
 Score: -0.042063
 Score: -0.042567
 Score: -0.042985
 Score: -0.043199
 Score: -0.043255
 Score: -0.043418

594 Aaron Shwartz is dead - Reddit, we did it!
 305 How do you rank the 32 starting Quarterbacks?
 506 RIP Aaron Swartz
 357 New Game Releases – January 2013
 85 PETER DINKLAGE WINS THE GOLDEN GLOBE!!!
 350 New Video Game Releases – January 2013
 300 Over 1/3 of starting QB's are in their rookie or sophomore season.
 610 [H] Skyrim, Arma 2 CO, CS:GO, Walking Dead, DmC Devil May Cry, Sid Meier's Civilization
 V - Gods and Kings, Arma X, FM 2013, Dishonored, Blackops 2 [W] TF Keys, Offers
 368 Miss Alabama Katherine Webb thinks being called 'Sexy' is derogatory.

356 New Game Releases – January 2013 - Q 103 - Albany's #1 Rock Station
 595 So is Bruce Banner Back?
 342 Post-Gazette.com - A Leno joke walked into 'Conan' - Pittsburgh Post-Gazette
 339 Pats hope to continue success without Gronkowski - NewsOK.com
 351 DMC - Devil May Cry's Sexy Character Reveal — Kat
 326 who let this guy host?
 349 92.9 WLMI Lansing's Greatest Hits - TV Tidbits - 'Mindy Project' Casting Shake-Up, 'Simpsons' Produce Judd Apatow Script - 92.9 WLMI Lansing's Greatest Hits
 190 The Island Packet - College Football
 379 WFMY News 2 - Naomi Campbell Mugged, George Lucas Engaged - Video - digtriad.com
 151 Husband in love with Selena Gomez
 286 Online activist Aaron Swartz commits suicide, relative says – CNN.com

Get position of relevant Documents in Ranking:

590 21
 591 264
 592 69
 593 60
 594 0
 595 10
 596 110
 597 568
 598 36
 599 47

Median of Ranking-Result: 53.5

In den Top 20 Ergebnissen sind 2 von 10 nach dem Testkorpus relevanten Dokumente vertreten. Einige der Artikel sind jedoch offenbar doch relevant fuer die Anfrage, aber nicht im Testkorpus als relevant markiert. Dies liegt an der vereinfachten Annahme, dass nur die bei der Query auftretenden Ergebnisse auch relevant sind, was offensichtlich nicht zutrifft. Im folgenden ein Beispiel.

```
In [21]: corpus.ix[506]["summary"]
```

```
Out[21]: "We are in recess until Monday in honor of the late Aaron Swartz, Reddit's Co-Founder."
```

1.3 Anfrage 3

```
In [22]: # Query: justin timberlake
         corpus[corpus["query"] == "justin_timberlake"]
```

```
Out[22]:
```

	source	query	score	\
690	news	justin_timberlake	1	
691	news	justin_timberlake	1	
692	news	justin_timberlake	1	
693	news	justin_timberlake	1	
694	reddit	justin_timberlake	19	
695	news	justin_timberlake	1	
696	news	justin_timberlake	1	
697	news	justin_timberlake	1	
698	news	justin_timberlake	1	

699 news justin_timberlake 1

headline \

690 Justin Timber...
691 99.5 Magic FM - From ABC News
692 Sun Herald - Highlights of Ho...
693 Tip(s) of the hat - StarTribune.com</s...
694 How come musicians often become well-respected...
695 Front - Highlights of Hollywo...
696 Ohio - Oscar Documentary Semi...
697 Product News - Panasonic Televisions to Offer ...
698 Forget about the Oscars, Justi...
699 WKBT - Celebrities with Janua...

summary

690 Get The Commercial Appeal on your iPhone, iPad...
691 There is a precedent for musicians who've ...
692 "Runner Runner";: An online gambling ...
693 Justin Timber...
694 Mark Wahlberg, Justin Timberlake, Will Smith, ...
695 "Runner Runner";: An online gambling ...
696 List is here I expect to have more to say late...
697 Music Star Justin <st...
698 ONLINE TRENDS FOR JAN 10, 2013. By Tim Swift,...
699 If January has you celebrating a birthday, you...

Ranking Resultate

Score: -0.045070
Score: -0.049264
Score: -0.055307
Score: -0.057209
Score: -0.057606
Score: -0.058397
Score: -0.060346
Score: -0.060892
Score: -0.062312
Score: -0.062392
Score: -0.062725
Score: -0.063486
Score: -0.064333
Score: -0.064550
Score: -0.065284
Score: -0.065646
Score: -0.065715
Score: -0.065915
Score: -0.066097
Score: -0.066185

321 Friday's TV talk shows - StarTribune.com

594 Aaron Schwartz is dead - Reddit, we did it!
 641 NEW YORK - Timberlake hints return to music in video - The Sun Herald - Celebrities
 305 How do you rank the 32 starting Quarterbacks?
 325 KTVO-TV 3 - Justin Bieber gets love at American Music Awards - The latest local news, weather and high ...
 320 HeraldNet.com - Premieres, finales abound on January TV schedule - HeraldNet.com - Entertainment
 762 Belfast, Northern Ireland, UK, World, News, Business, Entertainment, belfasttelegraph.co.uk - Independent Woman - Belfasttelegraph.co.uk
 264 Lance Armstrong Interview Could Be Worth Millions to Oprah Winfrey Network - Forbes
 691 99.5 Magic FM - From ABC News
 379 WFMY News 2 - Naomi Campbell Mugged, George Lucas Engaged - Video - digtriad.com
 695 Front - Highlights of Hollywood's 2013 film lineup - Movie News - Tri-CityHerald.com
 692 Sun Herald - Highlights of Hollywood's 2013 film lineup - The Sun Herald - Celebrities
 694 How come musicians often become well-respected actors, but never the other way round?
 122 Who won this trade?
 698 Forget about the Oscars, Justin Timberlake has the big the reveal today - baltimoresun.com
 349 92.9 WLMI Lansing's Greatest Hits - TV Tidbits - 'Mindy Project' Casting Shake-Up, 'Simpsons' Produce Judd Apatow Script - 92.9 WLMI Lansing's Greatest Hits
 496 NFL kicks fans in the face with Super Bowl XLV - We Never Stop Talking Baltimore Sports
 673 Lance Armstrong at ease about 'no-holds barred' interview with Oprah Winfrey - Dallas-Fort Worth Local Breaking ... 19 Wow.
 332 As a Pats fan I am going to do what so many college
 Get position of relevant Documents in Ranking:
 690 76
 691 8
 692 11
 693 95
 694 12
 695 10
 696 25
 697 23
 698 14
 699 449

Median of Ranking-Result: 18.5

In den Top 20 Ergebnissen sind 5 von 10 nach dem Testkorpus relevanten Dokumente vertreten. Hier gilt die selbe Beobachtung und Kritik an dem gewählten Evaluationsansatz, wie folgendes Beispiel zeigt (ebenfalls nicht im Testkorpus als relevant vertreten):

```
In [23]: corpus.ix[641]["summary"]
```

```
Out[23]: 'Timberlake hints <strong><strong>return</strong> <strong>to</strong> <strong>music</strong>
```

1.4 Anfrage 4

```
In [24]: # Query: devil may cry
         corpus[corpus["query"] == "devil_may_cry"]
```

```

Out[24]:
source      query      score \
350 news      devil_may_cry      1
351 news      devil_may_cry      1
352 news      devil_may_cry      1
353 reddit    devil_may_cry      960
354 news      devil_may_cry      1
355 reddit    devil_may_cry      8
356 news      devil_may_cry      1
357 news      devil_may_cry      1
358 reddit    devil_may_cry      5
359 reddit    devil_may_cry      27
610 reddit    devil_may_cry      2
611 news      devil_may_cry      1
612 reddit    devil_may_cry      584
613 news      devil_may_cry      1
614 reddit    devil_may_cry      15
615 reddit    devil_may_cry      4
616 reddit    devil_may_cry      29
617 news      devil_may_cry      1
618 news      devil_may_cry      1
619 reddit    devil_may_cry      3

                                         headline \
350 New Video Game Releases January 2013
351 DMC - <strong><strong>Devil</strong> <strong><strong>...
352 Too Human Review - MLive.com
353 Steam Summer Camp Sale Day 2: Total War Shogun...
354 <strong>Chicago Tribune</strong> - Monday melt...
355 WSIB: If I Liked the Devil May Cry Demo
356 New Game Releases January 2013 - Q 103 - Alb...
357 New Game Releases January 2013
358 [USA] [H] XCOM: Enemy Unknown, Dishonored, Bor...
359 IAmA music producer for Sony Music Japan and h...
610 [H] Skyrim, Arma 2 CO, CS:GO, Walking Dead, Dm...
611 <strong>LEO Weekly</strong> - Video TapeWorm -...
612 This Week in Gaming: October 16 - 22
613 <strong>ContraCostaTimes.com</strong> - A look...
614 /r/Gaming, what video game do you think revolu...
615 [USA] [H] Kinect, list (360), Dark Souls, God ...
616 IAmA music producer for Sony Music Japan and h...
617 New screenshots released for <strong><strong>D...
618 <strong>Austin's source for local news</strong>...
619 [USA] [H] Borderlands 2, Darksiders 2, Dead or...

                                         summary
350 <strong><strong>Devil</strong> <strong><strong>...
351 <strong><strong>Devil</strong> <strong><strong>...
352 Sure Too Human could have been a button masher...

```

353 Me again. \n \nhttp://store.steampowered.com...
 354 Taking a cue from "Silent Hill," the...
 355 I just installed the Devil May Cry Demo and ...
 356 Devil ...
 357 Devil ...
 358 Games are in **pristine** condition, and come ...
 359 This AMA is for my gf. She's worked at Sony Mu...
 610 * [ARMA II: Combined Operations](http://store...
 611 The aftermath may be the most...
 612 **Mobile Version**](https://sites.google.com/...
 613 The fun starts immediately in January with the...
 614 Just curious what everyone's opinions are on t...
 615 **Haves**\n\nXbox 360:* \n\n* ~~Kinect device...
 616 This AMA is for my gf. She's worked at Sony Mu...
 617 The Kids are Alright in Dungeon Defenders. Sl...
 618 I'm wondering if it's going to be kind...
 619 **Games I have**\n\n\n360:\n\n\n* borderlands ...

Ranking Resultate

Score: 0.000917
 Score: -0.001187
 Score: -0.003866
 Score: -0.004415
 Score: -0.004555
 Score: -0.005259
 Score: -0.005286
 Score: -0.008320
 Score: -0.009146
 Score: -0.010330
 Score: -0.011339
 Score: -0.011571
 Score: -0.011612
 Score: -0.011972
 Score: -0.012263
 Score: -0.012410
 Score: -0.012923
 Score: -0.012954
 Score: -0.012957
 Score: -0.013192

352 Too Human Review - MLive.com
 351 DMC - Devil May Cry's Sexy Character Reveal — Kat
 357 New Game Releases – January 2013
 76 Can we call progamers" something else?
 618 Austin's source for local news - Joystick takes an early look at Deadpool game - The Austin Daily Herald
 350 New Video Game Releases – January 2013
 355 WSIB: If I Liked the Devil May Cry Demo
 610 [H] Skyrim, Arma 2 CO, CS:GO, Walking Dead, DmC Devil May Cry, Sid Meier's Civilization

V - Gods and Kings, Arma X, FM 2013, Dishonored, Blackops 2 [W] TF Keys, Offers
 356 New Game Releases – January 2013 - Q 103 - Albany's #1 Rock Station
 775 PCMag.com - Web Mourns Internet Activist Aaron Swartz, 26 - News & Opinion - PCMag.com
 810 Any opinions on if the Wii U version of Aliens: Colonial Marines will have a Collector's Edition?
 193 Boston.com - Duped by the Due Date Mind Game - Ultra Sound Pregnancy - Boston.com
 103 A woman from my office was killed by her husband earlier this week.
 49 There's always some eccentric person in your city that you always seem to see, describe him/her/other.....
 574 Up till this day, I still regret seeing the movie Indiana Jones and the Kingdom of the Crystal Skull (the fourth movie).
 75 Hey Reddit.
 77 Hey fellow ents[7], I want to know if my brother is an ent.
 595 So is Bruce Banner Back?
 74 TIFU I read her diary.
 171 Game give-away.

Get position of relevant Documents in Ranking:

350 5
 351 1
 352 0
 353 147
 354 29
 355 6
 356 8
 357 2
 358 46
 359 126
 610 7
 611 86
 612 298
 613 26
 614 64
 615 152
 616 127
 617 25
 618 4
 619 214

Median of Ranking-Result: 27.5

In den Top 20 Ergebnissen sind 8 von 20 nach dem Testkorpus relevanten Dokumente vertreten. Hier faellt auf das weitere Testergebnisse thematisch ebenfalls Computerspiele thematisieren, jedoch nicht direkt das Spiel mit dem Titel "devil may cry". Im folgenden ein Beispiel:

```
In [25]: corpus.ix[171] ["summary"]
```

```
Out[25]: "Tired? Stressed? Too broke to buy that game everyone has been talking about for years?"
```