
Delivery 1 - Collocations

CSE4095 - Special Topics in Comp. Eng. III

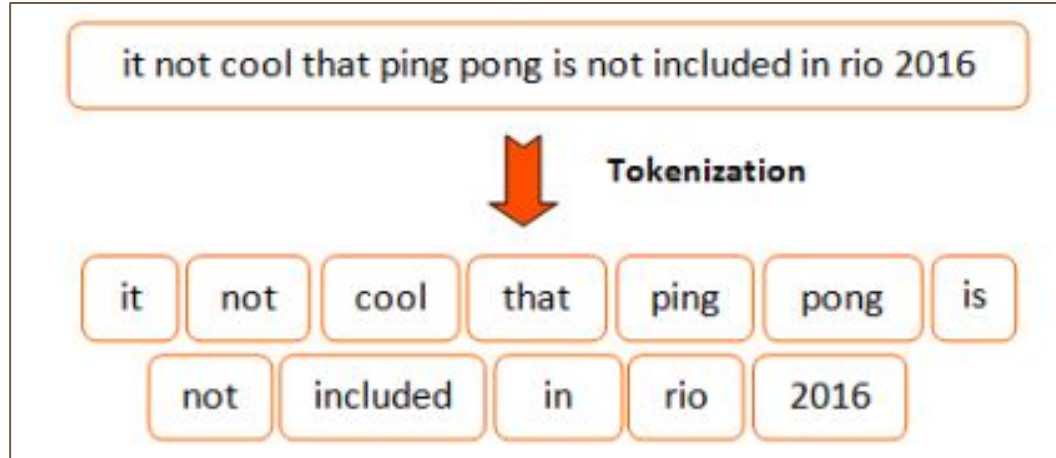
— Ali Reza Ibrahimzada
Muhammed Fatih Öztel
Basil Osama Hamid Elkhaila

Huzeyfe Ayaz
Yunus Ahmed Stahlschmidt
Sameeh Kunbargi

Content

- Data Cleaning and Preprocessing
- Frequency
- Mean and Variance
- Mutual Information
- Pearson's Chi-Square Test
- Dunning's Likelihood Ratio Test
- T-Test
- Text Summarization
- Contextualized Topic Modeling

Data Cleaning and Preprocessing



Moreover We Remove:

- Punctuations
- Digits
- Abbreviations
- etc.



Frequency

C(w1,w2)	w1	w2	TP
22186	sayılı	maddesi	NN
19283	bölge	adliye	NN
19280	maddesi	uyarınca	NN
13085	maddesi	ile	NN
13013	maddesi	gereğince	NN
13005	adliye	mahkemesi	NN
11092	sanık	hakkında	NN
11002	oy	birliğiyle	NN
10995	birliğiyle	karar	NN
10237	tarihinde	oy	NN
10024	temyiz	itirazlarının	NN
8794	sayılı	kanunun	NN
8530	yer	alan	NN
8498	ile	karar	NN
8352	vekili	tarafından	NN
8345	temyiz	itirazları	NN
8068	karar	sayılı	NN
7760	mahkemece	verilen	NN
7516	kanuna	uygun	NN
7075	sayılı	yasanın	NN

C(w1,w2,w3)	w1	w2	w3	TP
12997	bölge	adliye	mahkemesi	NNN
10995	oy	birliğiyle	karar	NNN
8634	sayılı	maddesi	uyarınca	NNN
7611	tarihinde	oy	birliğiyle	NNN
6115	verilen	hüküm	temyiz	NNN
5385	esas	karar	sayılı	NNN
5354	dosya	incelendi	gereği	NAN
5313	sayılı	maddesi	ile	NNN
5075	sayılı	maddesi	gereğince	NNN
4846	sayılı	kanunun	maddesi	NNN
4665	temyiz	itirazlarının	reddiyle	NNN
4489	dosya	incelenerek	gereği	NAN
4387	tarafklar	arasında	görülen	NNN
4324	sayılı	yasanın	maddesi	NNN
4318	maddesi	ile	sayılı	NNN
4308	mahkemece	verilen	hüküm	NNN
3902	tetkik	hakimi	tarafından	NNN
3870	adliye	mahkemesi	hukuk	NNN
3787	resmi	yayımlanarak	yürürlüğe	NNN
3785	tarafından	düzenlenen	rapor	NNN

Mean and Variance

s	d	w1	w2
0.30	+1.01	kasten	yaralama
0.42	+2.13	kasten	suçundan
0.30	-1.01	yaralama	kasten
0.25	+0.99	yaralama	suçundan
0.42	-2.13	suçundan	kasten
0.25	-0.99	suçundan	yaralama
0.00	+2.00	suçundan	nin
0.00	-2.00	nin	suçundan
0.00	+2.00	nin	türk
0.00	-2.00	türk	nin
0.31	+0.98	maddeleri	uyarınca
0.00	+3.00	maddeleri	lirası
0.31	-0.98	uyarınca	maddeleri
0.25	+2.07	uyarınca	lirası
0.20	+2.04	türk	adli
0.00	-3.00	lirası	maddeleri
0.25	-2.07	lirası	uyarınca
0.47	+1.14	lirası	adli
0.22	+2.00	lirası	para
0.20	-2.04	adli	türk

Mutual Information

I	w1	w2	w1 w2	Bigram
22.59	1	1	1	abana motorsan
22.59	1	1	1	abdurrahman karademir
22.59	1	1	1	abdusselam selmo
22.59	1	1	1	acizliğini örtbas
22.59	1	1	1	acıya gark
22.59	1	1	1	adaletine benzememektedir
22.59	1	1	1	ahkamına tevfikân
22.59	1	1	1	ailenle helalleş
22.59	1	1	1	aktiflerin azalmasına
22.59	1	1	1	akyiğit içtihatlı
22.59	1	1	1	aküleri teaşmetre
22.59	1	1	1	akıtılmasına katlanmakla
22.59	1	1	1	alacaklılarla uzlaşmaktan
22.59	1	1	1	alakadar olamamasını
22.59	1	1	1	alcool iode
22.59	1	1	1	aldatıyor musun
22.59	1	1	1	aleminyum folyoya
22.59	1	1	1	aleniliğin sonucuda
22.59	1	1	1	aleniyete bürüneceğinden
22.59	1	1	1	algılayan sensörlerin

Pearson's Chi-Square Test

X^2	C(w1, w2)	C(-w1, w2)	C(w1, -w2)	C(-w1, -w2)	Bigram
6317764.0	6	0	0	6317764	abalı rk
6317764.0	1	0	0	6317764	abana motorsan
6317764.0	1	0	0	6317764	abdurrahman karademir
6317764.0	1	0	0	6317764	abdusselam selmo
6317764.0	3	0	0	6317764	abercrombie fitch
6317764.0	2	0	0	6317764	aboneyi besleyen
6317764.0	2	0	0	6317764	absolute excellence
6317764.0	1	0	0	6317764	acizliğini örtbas
6317764.0	1	0	0	6317764	acıya gark
6317764.0	1	0	0	6317764	adaletine benzememektedir
6317764.0	3	0	0	6317764	ahi evran
6317764.0	1	0	0	6317764	ahkamına tevfikan
6317764.0	1	0	0	6317764	ailenle helalles
6317764.0	2	0	0	6317764	akciğerde kontüzyona
6317764.0	1	0	0	6317764	aktiflerin azalmasına
6317764.0	3	0	0	6317764	akvaryumlar livaryumlar
6317764.0	1	0	0	6317764	akyiğit içtihatlı
6317764.0	1	0	0	6317764	aküleri teaşmetre
6317764.0	2	0	0	6317764	akülerini sökmeye
6317764.0	1	0	0	6317764	akıtılmasına katlanmakla

$$X^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Dunning's Likelihood Ratio Test

```

Return boost::binom_pmf(x, n, p)

Ratio C(w1) C(w2) C(w1w2) w1 w2
1479.52 130 78 69 sonuçlarını algılama
1474.95 85 133 70 işlenmesindeki özellikler
1474.94 81 97 67 eli ürünü
1474.55 14406 121 121 oy çokluğuyla
1474.54 226 173 82 kötü niyetli
1473.53 1449 1093 136 ikinci kez
1472.12 194 1199 101 gönderilmekle okundu
1471.45 1238 1973 146 mahallinde düzeltilmesi
1470.33 734 5898 160 vaki düzeltme
1469.73 386 2609 125 öldüğünün anlaşılması
1469.55 388 147 85 elektronik haberleşme
1468.94 9757 328 148 katılan mağdure
1468.83 21611 3030 299 olduğu halde
1467.97 894 12827 193 bunun üzerine
1467.91 3801 26526 340 nedenlerle davalı
1467.72 1351 91 88 görevli görmez
1466.63 2564 650 136 davaya cevap
1466.58 12158 2803 251 tl vekalet
1466.24 15250 192 140 davanın açılmamış
1465.62 332 32501 182 kapsayan dosya
    
```

$$\log \lambda = \log \frac{L(H_1)}{L(H_2)}$$

$$L(H_1) = b(c_{12}; c_1, p) b(c_2 - c_{12}; N - c_1, p)$$

$$L(H_2) = b(c_{12}; c_1, p_1) b(c_2 - c_{12}; N - c_1, p_2)$$

$$b(k; n, x) = \binom{n}{k} x^k (1 - x)^{(n-k)}$$

T-Test

t	C(w1)	C(w2)	C(w1,w2)	w1	w2
164.67	87070	27890	27879	karar	verildi
142.49	106048	57284	22186	sayılı	maddesi
138.41	20785	19318	19283	bölge	adliye
136.91	57284	29714	19280	maddesi	uyarınca
113.49	19318	20676	13005	adliye	mahkemesi
112.6	57284	18553	13013	maddesi	gereğince
110.26	196076	106048	18143	ve	sayılı
107.66	95234	12011	11951	temyiz	edilmekle
107.54	57284	86368	13085	maddesi	ile
104.65	14406	11003	11002	oy	birliğiyle
104.12	34568	23163	11092	sanık	hakkında
103.41	11003	87070	10995	birliğiyle	karar
101.64	11974	20769	10409	basit	yargılama
100.5	29920	14406	10237	tarihinde	oy
98.58	95234	10259	10024	temyiz	itirazlarının
98.0	9671	21074	9669	görüşülüp	düşünüldü
92.19	20769	11087	8571	yargılama	usulü
92.1	16579	9003	8530	yer	alan
91.69	9703	196076	9000	usul	ve
91.56	106048	12380	8794	sayılı	kanunun

$$t = \frac{\bar{x} - \mu}{\sqrt{\frac{s^2}{N}}}$$

Text Summarization

Kasten basit yaralama suçundan sanık ...'nin, 5237 sayılı Türk Ceza Kanunu'nun 86/2, 86/3-e, 29, 62 ve 52/2. maddeleri uyarınca 2.240,00 Türk Lirası adli para cezası ile cezalandırılmasına, 5271 sayılı Ceza Muhakemesi Kanunu'nun 231/5. maddesi gereğince hükmün açıklanmasının geri bırakılmasına dair İstanbul Anadolu 40. Asliye Ceza Mahkemesinin 11.02.2020 tarihli ve 2019/750 Esas, 2020/94 Karar sayılı kararına karşı katılan sanık Ahmet Demir ve Cumhuriyet savcısı tarafından yapılan itirazın reddine ilişkin mercii İstanbul Anadolu 12. Ağır Ceza Mahkemesinin 14.07.2020 tarihli ve 2020/1000 değişik iş sayılı kararına karşı Adalet Bakanlığının 13.11.2020 tarihli ve 2020/14855 sayılı yazısıyla kanun yararına bozma isteminde bulunulduğundan bu işe ait dava dosyası Yargıtay Cumhuriyet Başsavcılığının 15.12.2020 tarihli ve 2020/103892 sayılı tebliğnamesi ile Dairemize gönderilmekle incelendi. Mezkur ihbarnamede; Dosya kapsamına göre, mercii İstanbul Anadolu 12. Ağır Ceza Mahkemesince yalnızca şekil şartları yönünden yapılan inceleme sonucunda itirazın reddine karar verilmiş ise de; Yargıtay Ceza Genel Kurulunun 22.01.2013 tarihli ve 2012/10-534 Esas, 2013/15 Karar sayılı ilamıyla itiraz merciinin sadece şekli olarak değil, hem maddî olay, hem de hukukî yönden inceleme yapabileceği yönündeki kararı nazara alındığında, itiraz mercii tarafından katılan sanık Ahmet Demir ve Cumhuriyet savcısının esasa yönelik itirazları konusunda inceleme yapılarak sonucuna göre bir karar verilmesi gerektiği gözetilmeden, sadece hükmün açıklanmasının geri bırakılmasına ilişkin koşulların var olup olmadığı yönünden değerlendirme yapılarak itirazın reddine karar verilmesinde isabet görülmediğinden bahisle, 5271 sayılı CMK'nin 309. maddesi gereğince anılan kararın bozulması lüzumunun ihbar olunduğu anlaşıldı. Gereği görüşülüp düşünüldü: Yargıtay Ceza Genel Kurulunun 22.01.2013 gün ve 2012/534 Esas, 2013/15 Karar sayılı kararında da belirtildiği üzere, hükmün açıklanmasının geri bırakılmasına dair karara karşı yapılan itiraz üzerine, merciin incelemesini sadece şekli olarak değil, 5271 sayılı Kanun'un 267-271. maddeleri uyarınca hem maddî olay hem de hukukî yönden incelenmesi gerektiğinden, itiraz mercii tarafından sanıkların esasa yönelik itirazları konusunda inceleme yapılarak sonucuna göre bir karar verilmesi gerektiği gözetilmeden, sadece hükmün açıklanmasının geri bırakılmasına ilişkin koşulların var olup olmadığı yönünden değerlendirme yapılarak itirazın reddine karar verilmesinde isabet bulunmamaktadır. Bu nedenle, Adalet Bakanlığının kanun yararına bozma isteyen yazısına dayanan tebliğnamede ileri sürülen düşünce yerinde görüldüğünden; İstanbul Anadolu 12. Ağır Ceza Mahkemesinin 14.07.2020 tarihli ve 2020/1000 değişik iş sayılı kararının 5271 sayılı CMK'nin 309/4. maddesi gereğince kanun yararına BOZULMASINA, müteakip işlemlerin mahallinde yerine getirilmesine, dosyanın mahalline gönderilmek üzere Yargıtay Cumhuriyet Başsavcılığına TEVDİİNE, 04.01.2021 gününde oy birliğiyle karar verildi.

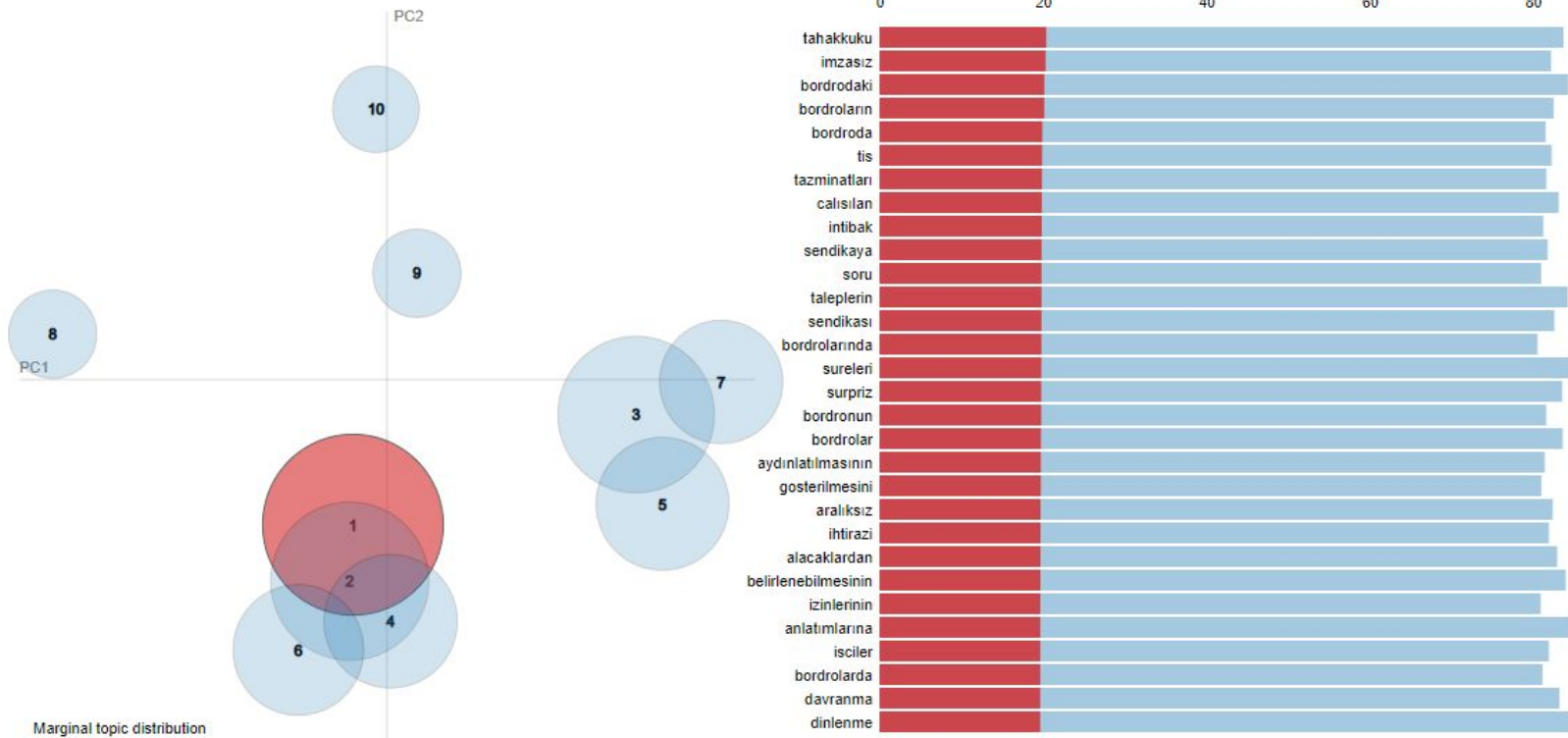
Text Summarization

Taraflar arasında görülen ve yukarıda açıklanan davada yapılan yargılama sonunda Ankara 20. Asliye Hukuk Mahkemesinin 15.12.2016 tarihli ve 2014/286 Esas, 2016/458 Karar sayılı kararıyla davanın karar verilmesine yer olmadığına karar verilmiş, Mahkeme hükmüne karşı davacı vekili tarafından istinaf yoluna başvurulması üzerine Ankara Bölge Adliye Mahkemesi 16. Hukuk Dairesince istinaf başvurusunun esastan reddine şeklinde hüküm kurulmuş olup, bu kez davacı vekilinin Bölge Adliye Mahkemesi kararını temyizi üzerine Dairece dosya incelendi, gereği düşünüldü: Tarafların karşılıklı iddia ve savunmalarına, dayandıkları belgelere, yapılan yargılama ve uyuşmazlık bakımından uygulanması gereken hukuk kuralları ile temyiz olunan Bölge Adliye Mahkemesi kararında yazılı gerekçelere, hukuki ilişkinin nitelendirmesine göre temyiz itirazları yerinde görülmediğinden, temyiz itirazlarının reddiyle usul ve kanuna uygun Bölge Adliye Mahkemesi kararının 6100 sayılı HMK'nin 370. maddesi uyarınca ONANMASINA, 31,40 TL peşin harcın onama harcına mahsubu ile kalan 27,90 TL'nin temyiz eden davacıdan alınmasına, 11.01.2021 tarihinde oy birliğiyle karar verildi

Text Summarization

- İstanbul Anadolu 12. Ağır Ceza Mahkemesi'nin 'kasten basit yaralama' suçundan sanık Ahmet Demir ve Cumhuriyet savcısı tarafından yapılan itirazın reddine ilişkin kararın bozulması lüzumunun ihbar edildiği bildirildi.
- Ankara Bölge Adliye Mahkemesi 16 Hukuk Dairesi'nde görülen ve 31-40 TL peşin harcının onama harcı ile kalan 27,90 TL'nin oy birliğiyle alınmasına karar verdi.

Contextualized Topic Modeling



THANK YOU FOR YOUR ATTENTION!!!

A Work By Members of Group 1