Batuhan Hökelek

Final Project Report

CSSM530: Automated Text Analytics for Social Sciences

Topic Modelling of Threats Concerning Refugee TikTokers in Turkey: GSDMM (Gibbs Sampling Dirichlet Multinomial Mixture), Latent Dirichlet Allocation (LDA) and LSI (Latent Semantic Indexing) for Turkish Anti-Refugee Tweets.

Problem Statement

TikTok is a video-based social-networking platform that is used to create various diversities of content such as lip-synching, comedy, dance, and talent videos usually made in short format. The platform particularly comes into prominence with its opportunities to include a broad range of user-profiles from K-pop lover teenagers to minorities and refugees in Turkey. TikTok offers a safe space for under-represented communities which are suppressed and despised to express their identity on other social media, to be more visible, and introduce their identity from their own hands. Analyzing outsider opinions about TikTok's content and user diversity on Twitter which has a more educated user base, makes it suitable to get an idea of attitudes toward under-represented groups of people who do not come from refined social strata but constitute a considerable proportion of Turkey's population. For example, starting from the 12th of April**, İstemiyorum** was a trending topic on Twitter mentioning that no more refugees are welcomed in Turkey. There were numerous tweets about refugees secretly filming non-consensual videos of women and minors and posting them on TikTok. This study is aimed to understand the type of threats expressed in tweets as a response to refugee videos on TikTok. Turkish tweets about TikTok and TikTok users were analyzed. For this task, 11634 tweets from this date to the 1st of June were extracted to spot refugee-related topics that might give insight into what type of threats are dominantly elicited from refugee videos on TikTok which are expected to be yielded from the clusters of topics.

1. Literature Review:

Threats about Refugee Migration:

Threat is an important concept for the prediction of attitudes toward refugees and migration. Integrated threat theory introduces several different types of threat: threat resulting from the concerns of cultural differences which is called the *symbolic threat*, specific negative aspects resulting from immigration such as financial strains which is a *realistic threat*, and *intergroup anxiety* which is manifested in expectations about conflict and negative interactions with out-group members. However, more recent research has expanded the scope of threats and added concerns about safety, infectious diseases, concerns that threaten the social coordination, personal freedom, and rights, and concerns regarding the distinctiveness of own group in comparison with out-groups. Several studies investigating refugee migration have approached the problem from only one threat dimension pointing out realistic, symbolic, safety, and/or health concerns (e.g., Louis, Duck, Terry, Schuller, & Lalonde, 2007; Yitmen & Verkuyten, 2018) and very few studies to date have explicitly distinguished between different threat types in the context of refugee migration. (e.g., Abeywickrama, Laham, & Crone, 2018; Tartakovsky & Walsh, 2016). A study by Lendman et. al. has mainly used a combination of qualitative and quantitative methods to identify and test the type of concerns and threats felt in the receiving society which was Germany in their case. This project aims to use topic modeling to extract hidden patterns of topics in the tweets related to such concerns and shed light on the reception of TikTok as a venue representative of refugees and sub-cultures in Turkey.

2. Data

Unlike previous research that mostly relies on qualitative analysis of certain threat types and their emotional reception in society, this project utilizes topic modeling to identify such threat types using social media data. Twitter is a text-based social media platform that is particularly purposed to share opinions, attitudes, and emotions and has totally different user clusters in terms of demographics and content tastes compared to TikTok. Therefore, Twitter is a suitable venue for identifying threats and emotional responses to the refugee video trends that started on TikTok.

Twitter API filtered with the keyword list below between 11th of April 2022 and 1st of May 2022 was scraped using the SNS TwitterScraper module. Since the scraper did not include tweets with different suffixes of the Turkish language, the keyword list was expanded with different suffix variations fed into it. Since some words such as “tiktok” were not in Turkish, only 11634 tweets labeled as “TR” were filtered out of 27000 tweets. The final dataset including tweets related to TikTok was acquired in a .json file and saved into a .csv (comma-separated value) file.

```
keywords = ["tiktok","tiktok'u","tiktoka","tiktok'a","tiktokun","tiktok'un","tiktokta","tiktok'ta","tiktokda",
            "tiktok'da","tiktoktan","tiktok'tan","tiktokdan",

           "tiktoktaki","tiktok'taki","tiktok'daki","tiktokdaki","tiktoktakine","tiktoktakini","tiktoktakinde",
            "tiktoktakinden",

           "tiktoktakiler","tiktokdakiler","tiktoktakilere","tiktokdakilere","tiktoktakilerde","tiktoktakilerden",
           "tiktoktakilerin","tiktokdakilerin",

           "tiktokcu","tiktokçu","tiktokcuya","tiktokçuya","tiktokcuyu","tiktokçuyu"]
```

3. Pre-Processing:

After acquiring tweets data, the following steps were applied to convert the text into the appropriate format for topic modeling.

1) The dataset was cleaned by removing capital letters, tagged users, punctuation, numbers, hyperlinks, and emojis and converted into the lists of tokens by applying regular expressions. Stop words were also dropped from the tweets by importing Turkish stop words from nltk.corpus Python package.

2) The words in the dataset were normalized which means that misspelled words are corrected to their standard formats. This step was achieved by using Zemberek library in Python.

3) Finally, lemmatization was applied to the words in the dataset by using Zemberek library in Python. Lemmatization refers to a process in which words with the same base are reduced to their basic forms and taken as a single item (e.g. words “spoken” and “speaking” are both counted as their common lemma of “speak”).

Methodology / Topic Modelling:

Topic modeling is an unsupervised machine learning technique that is extensively used to find hidden patterns in data to spot clusters of topics. It is widely used in many spheres to give insights into the nature of discussions concerning the public and societal problems. Choi and Jang investigated public opinions of acceptance for asylum seers in Korean society. They used topic modelling to identify the main ideas and the context of comments under the videos related to refugees on social media (2010). In another study by Heidenreich et. al., LDA was applied for a similar analysis of national media discourses related to refugees in several European countries (2019). LDA (Latent Dirichlet Allocation) is an example of probabilistic topic modeling. LDA assumes that each document or tweet includes a mixture of topics that consist of a bunch of words and outputs the varying probabilities of each topic’s contribution to the document. On the other hand, GSDMM (Gibbs Sampling Dirichlet Multinomial Mixture) is another method deriving from the LDA technique that assumes that one subject is matched with one document. The words within a document are generated using the same topic rather than a mixture of topics as in LDA. GSDMM is a technique that outperforms LDA for short text topic modeling because of its advantages such as automatic identification of the number of clusters and its fast convergence rate. Another problem with short text that can be eliminated by GSDMM is that short text is represented by sparse vectors of high dimensional space (Akritidis et. al., 2020). GSDMM also performs well in capturing the representative words of topic clusters. Latent Semantic Indexing (LSI) is another method that differs from LDA in that it

does not focus on the relationships between documents making up a corpus. LDA, on the other hand, assumes that documents with a higher proportion of similar words and groups of words are about the same topic.

Topic Models:

GSDMM: The dataset that was applied lemmatization, was saved in a .csv file, read and inspected. We created n-grams for the GSDMM. MovieGroupProcess() algorithm hyperparameters were set to: alpha =0.01 1 and beta (eta) = 0.01. n – items = 30. The number of clusters was set to 7. This was the most optimal number we identified after experimenting with more clusters that outputted empty results and fewer clusters that yielded less relevant topics. The dominant words for each topic were visualized with WordCloud.


```
Number of documents per topic : [1996 1656 1887 1705 1534 1581 1275]

Most important clusters (by number of docs inside): [0 2 3 1 5 4 6]

Cluster 0 : [('tıkmak', 1179), ('bir', 554), ('demek', 427), ('olmak', 380), ('mi', 294), ('ben', 290), ('video', 256), ('tiktokçu', 226), ('var', 21
3), ('se', 204)]

Cluster 2 : [('tıkmak', 1132), ('bir', 418), ('video', 355), ('demek', 283), ('olmak', 271), ('ben', 247), ('tıktokçu', 174), ('se', 170), ('mi', 166),
('tiktokçu', 153)]

Cluster 3 : [('tıkmak', 862), ('bir', 425), ('demek', 339), ('olmak', 287), ('ben', 249), ('tiktokçu', 229), ('video', 208), ('tıktokun', 184), ('tıkto
kçu', 184), ('var', 176)]

Cluster 1 : [('tıkmak', 990), ('bir', 375), ('demek', 356), ('olmak', 289), ('video', 261), ('ben', 241), ('tıktokçu', 187), ('bilmek', 184), ('mi', 18
4), ('se', 147)]

Cluster 5 : [('tıkmak', 871), ('bir', 447), ('olmak', 371), ('demek', 369), ('video', 255), ('ben', 249), ('tiktokçu', 232), ('se', 190), ('tıktokçu',
166), ('kız', 131)]

Cluster 4 : [('tıkmak', 812), ('bir', 449), ('olmak', 314), ('demek', 306), ('video', 226), ('ben', 196), ('tıktokçu', 191), ('mi', 189), ('tiktokçu',
186), ('tıktokda', 175)]

Cluster 6 : [('tıkmak', 479), ('tıktokda', 410), ('bir', 280), ('video', 278), ('demek', 213), ('olmak', 193), ('var', 142), ('pay', 114), ('mi', 105),
('tiktokçu', 99)]
```

Figure 1. GSDMM Clusters

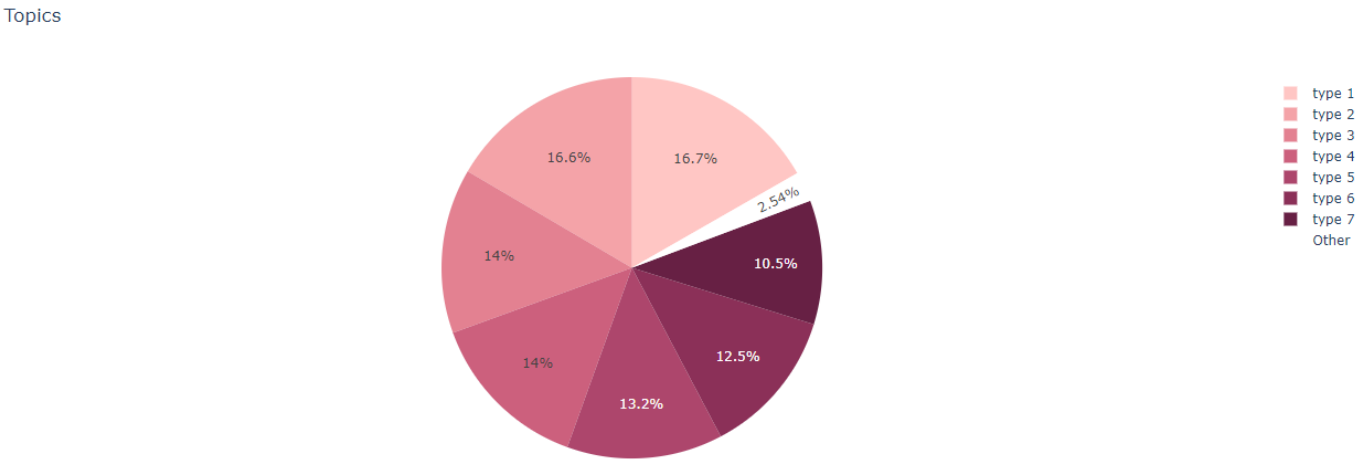

Figure 2. GSDMM Topics Chart

	Text	Topic	Rating	Lemma-text
0	chp zihniyet demek eğe böyle bir yapmak fanati...	type 4	NaN	[chp, zihniyet, demek, eğe, böyle, bir, yapmak...
1	usta çin tıktoku kul bütün veri çin almak göt ...	type 6	NaN	[usta, çin, tıktoku, kul, bütün, veri, çin, al...
2	iyi gecelemek tıktoku	type 2	NaN	[iyi_gecelemek, tıktoku]
3	tıktoku bırakmak iyi olmak nazar değmek hahshs...	type 6	NaN	[tıktoku, bırakmak, iyi, olmak, nazar, değmek,...
4	kızmak yer tıktoku temsil etmek ner fran loğ b...	type 6	NaN	[kızmak, yer, tıktoku, temsil, etmek, ner, fra...

Figure 3. GSDMM Topics Dataframe

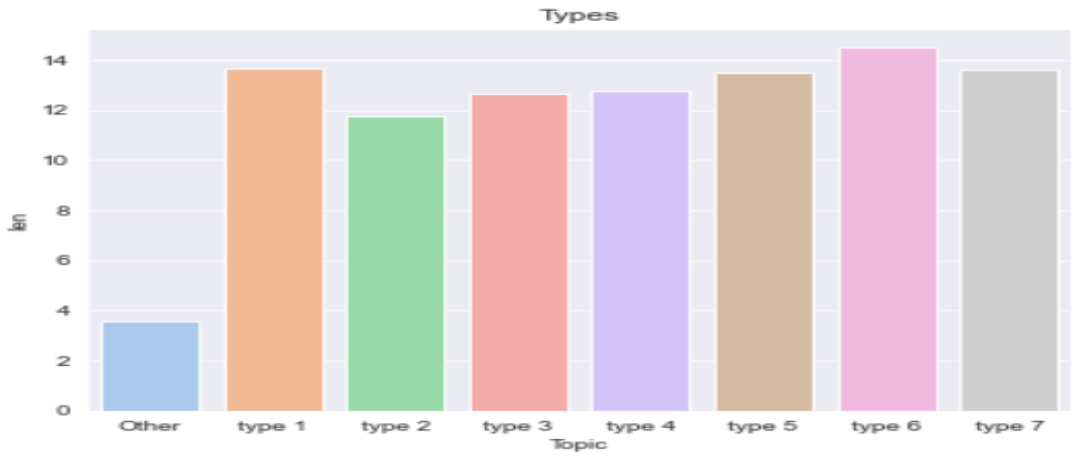

Figure 4: Dispersion of Topics with GDSMM

LDA, LDA Multicore and LSI:

The original LDA algorithm was implemented without applying the TF-IDF transformation on the lemmatized data by using Gensim LDA module. The Gensim dictionary was created by storing lemmatized tweets as tokens where each unique term was also assigned an index. The model was built by inputting the dictionary of tokens, the number of topics determined for the corpus, and the corpus itself. We printed the words with the highest rate of frequency for each topic. LDA method was used to create a document-topic matrix and a topic-word matrix storing the probabilities of words' distributions into topics. The topic-word matrix included 10 topics and words that best represent each topic. With the document-topic matrix, documents can be assigned to a variety of themes. PyLDA was used to plot the user interactive visualization of topics in relation to the text. Despite selecting the topic number with the highest coherence score, the first LDA model yielded overlapping topics. That's why we built LDAMulticore as well to accelerate model training. The LDAMulticore model was initiated, and the models with the best number of topics were created, we implemented a chart that displays the best coherence score for every possible number of topics. We then calculated the hyperparameters of the models with the number of topics that displayed the highest coherence score. These parameters were later fed into the LDAMulticore model with the optimal number of topics to get a coherence score of 0.68. For the LSI model, we found the number of topics with a chart that illustrates the best coherence scores. After setting the model on the highest coherence score, we acquired 2 topics as seen on the chart.

```
pprint(ldamodel.print_topics())
```

```
[(0,
  '0.005*"imek" + 0.005*"mamoğlu" + 0.004*"oğlan" + 0.004*"tıktokdaki" + '
  '0.003*"sanıyorum" + 0.003*"üst" + 0.003*"yarış" + 0.003*"kal" + '
  '0.003*"tiktokçuyu" + 0.003*"in"'),
 (1,
  '0.083*"tıktoku" + 0.015*"silmek" + 0.009*"bütün" + 0.006*"düşün" + '
  '0.005*"akım" + 0.004*"ukrayna" + 0.003*"kavga" + 0.003*"sizce" + '
  '0.003*"keşfetmek" + 0.003*"yeniden"'),
 (2,
  '0.014*"hal" + 0.005*"komik" + 0.005*"okumak" + 0.004*"veri" + '
  '0.004*"yarışmak" + 0.003*"pota" + 0.003*"şaka" + 0.002*"sponsor" + '
  '0.002*"oy" + 0.002*"kıyaslamak"'),
 (3,
  '0.015*"allah" + 0.010*"beyin" + 0.009*"göt" + 0.005*"sıkmak" + 0.005*"aşk" '
  '+ 0.003*"yaratmak" + 0.003*"fava" + 0.003*"çıkacak" + 0.002*"şaka_maka" + '
  '0.002*"koru"'),
 (4,
  '0.021*"artık" + 0.021*"takip" + 0.012*"biraz" + 0.011*"olucam" + '
  '0.010*"herkes_sevmek_tiktokçuyu" + 0.009*"edit" + 0.008*"kanka" + '
  '0.007*"saç" + 0.005*"ölürüm" + 0.005*"çıkı"'),
 (5,
  '0.030*"tiktokçu" + 0.014*"sapık" + 0.014*"ülke" + 0.014*"pakistan" + '
  '0.013*"afgan" + 0.010*"genç" + 0.010*"kadı" + 0.007*"türk" + 0.007*"taciz" '
  '+ 0.006*"suriye"'),
 (6,
  '0.022*"bura" + 0.015*"yine" + 0.014*"twitter" + 0.009*"çıktı" + '
  '0.006*"bayram" + 0.006*"başlamak" + 0.006*"nstagram" + 0.005*"uzak" + '
  '0.004*"konser" + 0.004*"prim"'),
 (7,
  '0.035*"kendi" + 0.021*"yaş" + 0.017*"herkes" + 0.015*"karşı" + '
  '0.008*"galiba" + 0.007*"tıktokçuya" + 0.006*"idol" + 0.006*"benzemek" + '
  '0.004*"tıktokçuyu" + 0.004*"doktor"'),
 (8,
  '0.022*"gerçek" + 0.009*"aç" + 0.008*"zekâ" + 0.007*"tiktokçu" + 0.007*"max" '
  '+ 0.005*"çakal" + 0.004*"tat" + 0.004*"parlamak" + 0.003*"hani" + '
  '0.003*"veli"'),
 (9,
  '0.035*"tıktokçu" + 0.033*"tiktokçu" + 0.027*"bir" + 0.027*"olmak" + '
  '0.024*"demek" + 0.014*"tıkmak" + 0.014*"ben" + 0.011*"mi" + 0.010*"se" + '
  '0.009*"var"')]
```

Figure 5. Topics for the first LDA model

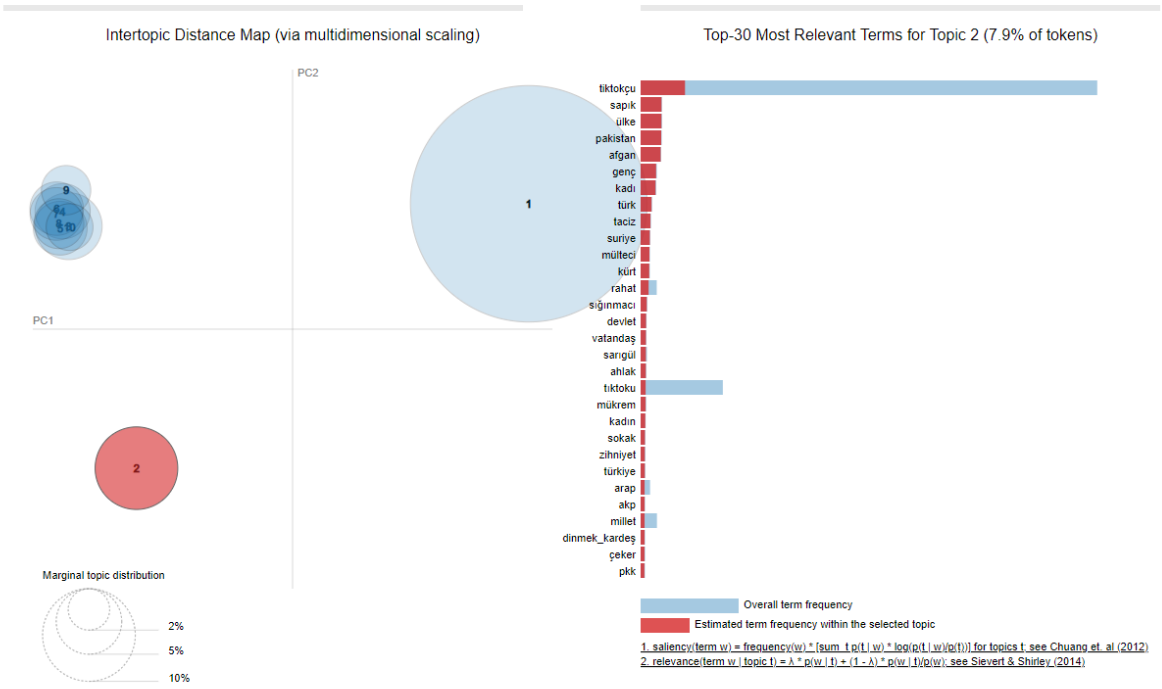

Figure 6. Dominant Topics for the first LDA model Coherence Score 0.54.

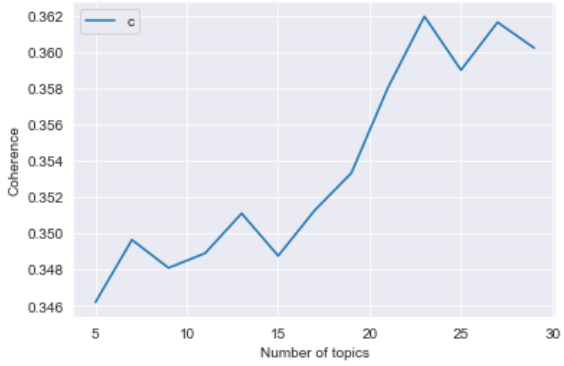

Figure 16. LDAMulticore coherence chart, optimal number of topics is 23.

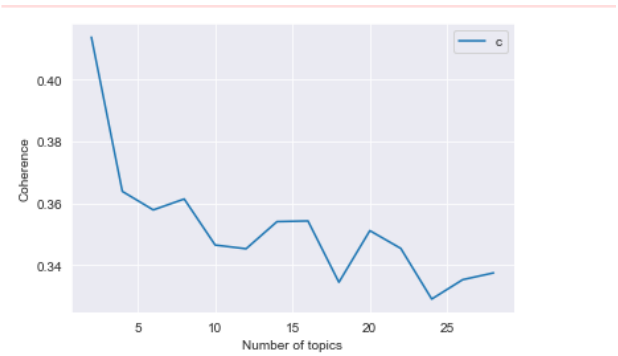

Figure 17. LSI coherence chart, optimal number of topics is 2.

	Document_No	Dominant_Topic	Topic_Perc_Contrib	Keywords	Text
0	0	9.0	0.5721	tıktokçu, tiktokçu, bir, olmak, demek, tıkmak,...	chp zihniyet demek eğe böyle bir yapmak fanati...
1	1	3.0	0.4211	allah, beyin, göt, sıkmak, aşk, yaratmak, fava...	usta çin tıktoku kul bütün veri çin almak göt ...
2	2	1.0	0.6991	tıktoku, silmek, bütün, düşün, akım, ukrayna, ...	iyi gecelemek tıktoku
3	3	9.0	0.3962	tıktokçu, tiktokçu, bir, olmak, demek, tıkmak,...	tıktoku bırakmak iyi olmak nazar değmek hahshs...
4	4	9.0	0.9139	tıktokçu, tiktokçu, bir, olmak, demek, tıkmak,...	kızmak yer tıktoku temsil etmek ner fran loğ b...
...	...	...	...	...	...
11629	11629	9.0	0.5339	tıktokçu, tiktokçu, bir, olmak, demek, tıkmak,...	tıktoku kadar seveceğim ak gelmek
11630	11630	9.0	0.4648	tıktokçu, tiktokçu, bir, olmak, demek, tıkmak,...	birlik tıkmak izleyebileceğim olmak tıktoku ku...
11631	11631	9.0	0.4789	tıktokçu, tiktokçu, bir, olmak, demek, tıkmak,...	amor tıktoku el geçirmek asi adam öldürmek yin...
11632	11632	9.0	0.7088	tıktokçu, tiktokçu, bir, olmak, demek, tıkmak,...	tıktoku ben mi açacak abi
11633	11633	9.0	0.6873	tıktokçu, tiktokçu, bir, olmak, demek, tıkmak,...	utanmak çok tıktoku aktif kul

Figure 18. Dominant Topics for the first LDA model

	Topic_Num	Topic_Perc_Contrib	Keywords	Text
0	0.0	0.4389	bir, tıkmak, demek, tiktokçu, video, ben, olma...	[evet, hemen, ahlak, edep, hayâ, nanç, bir, ka...
1	1.0	0.4448	olmak, bir, tıkmak, ben, demek, tıktokçu, tikt...	[tıkmak, geziyor, tam, süper, bir, jisung, vid...
2	2.0	0.4227	tıkmak, video, ben, tıktokçu, tiktokçu, mi, de...	[ertuğrul, gazi, dizi, pakistan, aşırı, popüle...
3	3.0	0.4236	tıkmak, tıktokçu, bir, demek, olmak, tiktokçu,...	[sude, fâni, aleyna, nisa, tibet, ağ, merak, e...
4	4.0	0.4161	tıkmak, video, tiktokçu, demek, olmak, ben, bi...	[influençerların, ağız, siçilmasina, izin, ver...
5	5.0	0.4178	tıkmak, olmak, tıktokçu, bir, tıktokda, ben, m...	[lik, dul, bunak, burcu, esmersoy, un, tıkmak,...
6	6.0	0.4154	tıktokçu, tıkmak, demek, se, video, olmak, bir...	[yeten, qorumaq, enez, arpad, uşak, qoyub, kaç...
7	7.0	0.4087	tıkmak, tiktokçu, tıktokçu, olmak, ben, bir, t...	[tam, olmak, burak, hoca, nternet, ev, fiyat, ...
8	8.0	0.4334	tıkmak, bir, tiktokçu, tıktokda, olmak, demek,...	[evet, tıkmak, engellemek, fenomen, aybüke, ça...
9	9.0	0.4201	olmak, tıkmak, bir, demek, tıktokçu, video, tı...	[dü, bir, dost, danışmak, kur, bir, niza, tıkt...
10	10.0	0.4199	ben, tıkmak, tıktokçu, bir, tıktokda, olmak, t...	[tıktoku, cox, sonradan, tanımak, biri, tıktok...
11	11.0	0.4083	tıkmak, mi, olmak, video, bir, demek, se, pay,...	[evrim, pota, nisa, just, poor, nbfç, berkan, ...
12	12.0	0.4200	bir, tıkmak, olmak, ben, demek, mi, video, tik...	[allah, eləməsin, döyəcəyim, son, qapıdı, amma...
13	13.0	0.4030	tıkmak, ben, tıktokçu, demek, bir, var, video,...	[dayıqizinin, hornby, kapan, təzə, açılış, büt...
14	14.0	0.4168	tıkmak, tiktokçu, ben, olmak, se, tıktokçu, tı...	[sahan, khrshd, bura, çelişki, görür, men, siz...
15	15.0	0.4185	tıkmak, tiktokçu, ben, olmak, bir, demek, tıkt...	[şimdi, genç, sanat, netflix, tıktokçu, instag...
16	16.0	0.4152	tıkmak, bir, mi, demek, video, olmak, tıktokçu...	[bir, sg, üstün, ırk, kafa, tr, de, dışarı, çı...
17	17.0	0.4433	olmak, tıkmak, bir, ben, demek, tıktokda, tıkt...	[ana, skdymi, suriyalıları, ran, galib, türkiy...
18	18.0	0.4196	tıkmak, tıktokçu, bir, video, olmak, demek, ti...	[lik, dul, bunak, burcu, esmersoy, un, tıkmak,...
19	19.0	0.4147	tıkmak, olmak, video, bir, demek, ben, tiktokç...	[ben, tıktokun, var, ay, ön, öğrenmek, üzer, b...
20	20.0	0.4172	tıkmak, olmak, tiktokçu, bir, mi, tıktokçu, de...	[lopes, yel, kadar, afgan, pek, gelmek, te, hi...
21	21.0	0.4119	tıkmak, tiktokçu, olmak, mi, bir, demek, video...	[sizfirilmiş, imek, öz, bang, demek, ihanet, e...
22	22.0	0.4252	tiktokçu, olmak, tıktokçu, tıkmak, bir, mi, vi...	[tik, tok, la, gelecek, bel, olmak, türkiye, g...

Figure 19. LDA Multicore Topics and Keywords

Discussion:

In this project, we performed several topic modeling algorithms GSDMM, LDA, and LSI on a subset of tweets related to refugee videos on TikTok after the spread of non-consensual videos and the negative reactions thereafter. Future work can further analyze the emotions expressed in the tweets by jointly using topic modeling with emotion detection so that richer insights can be concluded from the analysis of topics signaling specific concerns associated with their linked emotions. Several setbacks in the process of this project that, we noticed, could need further improvement are as follows:

1. The length and time interval of the dataset could be expanded. In this project, we included only a final dataset of 11634 tweets after the date that the refugee videos on TikTok came to the public's attention (11th of April with the trending hashtag "istemiyorum") . Replication of this study might include a time-series analysis of topics in combination with emotion detection before and after this date.
2. Turkish stop words downloaded from the NLTK library didn't suffice for the elimination of all stop words in the pre-processing phase of our study which led to the algorithms (GDSSM and LDA) building their clusters containing these missing sets of stop words.
3. Parameters we fed into GDSSM could be more carefully inspected and the model could be re-tested with new parameters.
4. In this project, we tuned the parameters of the LDA model using GridSearchCV() and looked at evaluation scores, and figured out the most optimal topic number, the created output was pathological overlapping of topics that need to be more equally dispersed (as shown in figure 6). Further research is needed to examine the clustering in more carefully.
5. LDAMulticore topics did better than the normal LDA model in terms of equal dispersion of topics. The elbow method was applied to decide on the number of topics and the hyperparameters were tuned. However, the coherence score of .50 needs improvement by a more careful tuning of the algorithm.

REFERENCES

Abeywickrama, R. S., Laham, S. M., & Crone, D. L. (2018). Immigration and receiving communities: The utility of threats and emotions in predicting action tendencies toward refugees, asylum-seekers and economic migrants. *Journal of Social Issues*, *74*(4), 756-773. doi:10.1111/josi.12297

Choi, S., & Jang, S. Y. (2019). An Online Opinion Analysis on Refugee Acceptance Using Topic Modeling. *Asian Journal for Public Opinion Research*, *7*(3), 169 198. https://doi.org/10.15206/ajpor.2019.7.3.169

Heidenreich, T., Lind, F., Eberl, J.-M., & Boomgaarden, H. G. (2019). Media Framing Dynamics of the 'European Refugee Crisis': A Comparative Topic Modelling Approach. *Journal of Refugee Studies*, *32*(Special_Issue_1), i172–i182. https://doi.org/10.1093/jrs/fez025

Landmann, H., Gaschler, R., & Rohmann, A. (2019). What is threatening about refugees? Identifying different types of threat and their association with emotional responses and attitudes towards refugee migration. *European Journal of Social Psychology*, *49*(7), 1401–1420. https://doi.org/10.1002/ejsp.2593

Louis, W. R., Duck, J. M., Terry, D. J., Schuller, R. A., & Lalonde, R. N. (2007). Why do citizens want to keep refugees out? Threats, fairness, and hostile norms in the treatment of asylum seekers. *European Journal of Social Psychology*, *37*(1), 53–73. https://doi.org/10.1002/ejsp.329

L. Akritidis, M. Alamaniotis, A. Fevgas and P. Bozanis, "Confronting Sparseness and High Dimensionality in Short Text Clustering via Feature Vector Projections," *2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI)*, 2020, pp. 813-820, doi: 10.1109/ICTAI50040.2020.00129.

Tartakovsky, E., & Walsh, S. D. (2016). Testing a new theoretical model for attitudes toward immigrants: The case of social workers' attitudes toward asylum seekers in Israel. *Journal of Cross-Cultural Psychology*, *47*(1), 72-96. doi:10.1177/0022022115613

Yitmen, Ş., & Verkuyten, M. (2018). Positive and negative behavioral intentions towards refugees in Turkey: The roles of national identification, threat, and humanitarian concern. *Journal of Community & Applied Social Psychology*, *28*(4), 230- 243. doi:10.1002/casp.2354