HOW DID PEOPLE REACT TO ISTANBUL CONVENTION?

A Sentiment Analysis of Tweets Regarding Istanbul Convention and How Topics Change Among Tweets with Positive Sentiment and with Negative Sentiment

Istanbul Convention, the Council of Europe Convention on Preventing and Combating Violence Against Women and Domestic Violence, is the first European Convention that targets particularly violence/domestic violence against women and young girls. The Convention aims to protect women from all forms of violence, prevent, prosecute, and eliminate violence against women and domestic violence (Council Of Europe, 2021). It contributes to eliminating all forms of discrimination against women and promoting fundamental equality between women and men through women's empowerment. It helps to develop a comprehensive framework, policies, and measures to protect and assist victims of violence. Adopting a holistic approach to eliminating violence against women and domestic violence provides support and assistance to organizations and law enforcement agencies.

Turkey was the first country to sign the Istanbul Convention on May 11, 2011, and ratify it in its parliament on November 24, 2011. The Convention was opened for signature at the 121st meeting of the Committee of Ministers of the Council of Europe held in Istanbul on May 11, 2011. That is why the Convention is known as the "Istanbul Convention". As of July 2020, it has been signed by 45 countries and the European Union. Although Azerbaijan and Russia are members of the Council of Europe, they did not sign the Convention. Although many countries signed the Convention, there was a trend for non-ratification of the Convention in Europe (Council of Europe, 2023). While some countries completed the ratification process even though the strong anti-gender movements in the West (e.g., France and Germany) and Central-Eastern European region (e.g., Croatia, and Poland), in the other countries (Bulgaria, the Czech Republic, Hungary, Latvia, Lithuania, Slovakia in Central-Eastern and only United Kingdom in the West), those anti-gender movements used as a legitimization for non-ratification of the Convention (Balogh, 2020). Despite being the early adopters of the Convention, as a result of the presidential decision published in the Official Gazette on March 20, 2021, Turkey has officially withdrawn from the agreement on July 1, 2021, as a result of Recep Tayyip Erdoğan's decision. The fact that the president of Turkey was the center of most criticism, this debate has turned into a political one.

Problem Statement: The withdrawal of Turkey from the Istanbul Convention has generated diverse sentiments among the public, ranging from support to criticism and everything in between. Understanding the sentiment expressed in tweets regarding the withdrawal is essential for gauging public opinion and informing relevant stakeholders about the implications of the decision.

Data & Methods

Data Collection

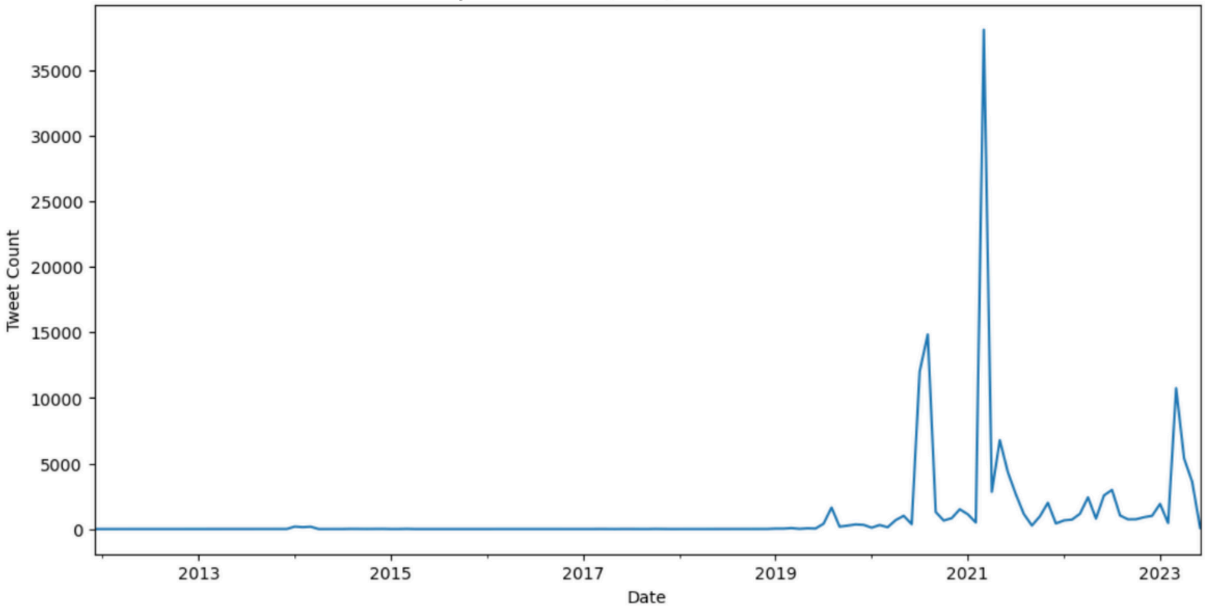
The primary data source for this project is tweets collected using SNS scraping tool. In addition to that some part of the data set was collected from ERC-funded Politus project[1], which aims to assess the commercial viability of proprietary natural language processing (NLP) and machine learning (ML) technology for AI-based data-driven fair social policymaking. The dataset also includes tweets including Istanbul Convention keyword, with some portion of the dataset annotated based on sentiment to train the sentiment analysis model.

---

[1] https://ccss.ku.edu.tr/tcss-2023-2__trashed/politus/

The time period has been selected as one year from the publishment of the Turkish president's, Recep Tayyip Erdoğan, decision to withdraw from the Istanbul convention in the Official Gazette on March 20, 2021. The time period has been selected as one year since Turkey officially withdrew from the contract on July 1, 2021, and tweets distribution is mainly concentrated on the first one year (Graph 1.). There are 48,692 tweets in the data set.

Graph 1. Tweet Frequency Over Time



Methodology

In this project, a Python-based data analysis pipeline has been employed. Sentiment analysis using natural language processing techniques and supervised-machine learning based prediction models were performed. To be able to implement SML, 3,000 tweets were randomly selected for the manual annotation and then used to predict the sentiment of the whole data set. 5 supervised machine learning algorithms were used in this study, which are Multinomial Naive Bayes, K-nearest Neighbors, Random Forest, Logistic Regression, and Support Vector Machine. To decide which model to use, f1 scores have been calculated for each model.

Annotation

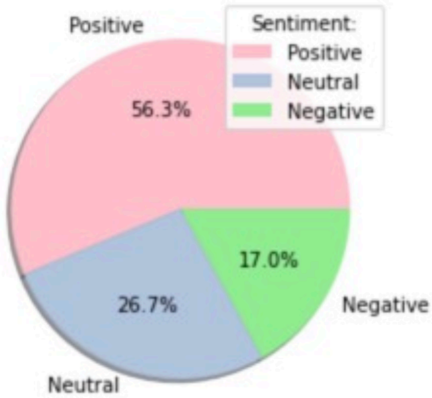
For the annotation part, 3,000 tweets were randomly selected and labeled manually by a single annotator as positive, neutral, and negative. Tweets that include positive sentiment towards the Istanbul Convention were labeled as **positive**. The other tweets that include negative opinions about the convention were labeled as **negative**. In addition to these, other tweets do not include any opinion regarding the convention were labeled as **neutral**. %56.3 of tweets were labeled as positive, %26.7 was negative, and the remaining %17 was labeled as neutral (Figure 1.). Some examples of annotated tweets and their translation are represented in Table 1.

Table 1. Some Examples from Annotated Tweet Data Set (TUR-ENG)

Id	Tweets	Labels
1,37321E+18	Sözleşmeye karşı çıkan herkesin bütün katliamlarda payı vardır #istanbulsoezlesmesi	positive
1,37318E+18	istanbul sözleşmesi bizimdir, bir aradayız ve asla vazgeçmeyeceğiz. #istanbulsözleşmesiyaşatır	positive
1,3732E+18	yemin ederim otobuse binmekten korkuyorum, sokakta yurumekten korkuyorum #istanbulsözleşmesi	positive
1,37321E+18	Artık ne benim güvenliğim, ne de bir başka kadının güvenliği yok. #istanbulsoezlesmesi	positive
1,37319E+18	İstanbul Sözleşmesi yaşatırdı, uygulansaydı. #istanbulSözleşmesi	positive
1,37321E+18	Bu ülkede her gün daha mutsuz uyanıyoruz. Her gün haklarımızı kaybederek ,susarak,susturularak.. Sözleşmenin kabulünün kimleri koruduğu değil kimlere zarar verdiği düşünülüyor. Katilleri koruma. Kadınları koru. #istanbulsözleşmesi	positive
1,37436E+18	Cumhurbaşkanı Erdoğan'ın gece yarısı bir karsrnameyle geri çektiği istanbul sözleşmesini Birleşmiş Milletler İnsan Hakları Ofisi, Türkiye'yi kadına yönelik şiddetle mücadele amaçlı uluslararası bir anlaşma olan İstanbul Sözleşmesi'nden çekilme kararını geri almaya çağırdı. https://t.co/jRzIWPethV	neutral
1,40491E+18	TBMM Kadına Yönelik Şiddetin Araştırılması Komisyonu, akademisyenleri dinledi https://t.co/cQbVxuc9Tw #kadınaşiddet #istanbulsözleşmesi https://t.co/JXFCcF2WH0	neutral
1,37361E+18	. Aslinda istanbul sözleşmesi neymiş izleyin.. Muhtesem anlatmis 👍 https://t.co/1UnJtUEzCv	neutral
1,37317E+18	#MerkezBankası başkanı değişti kimse konuşmasın diye de gündem #istanbulsoezlesmesi oldu. Bravo! Şimdi herkes içeriğini bile bilmeden yorum yapacak burası Türkiye 🤫	neutral
1,37328E+18	Bizim dinimiz kadına yeterince değer veriyor Eğer yasa yapılacaksa ölçü Kuran, Sünnet olmalı İslâmın emrettiği gibi olmalı Medeniyetsiz Batı değil #istanbulSözleşmesi	negative
1,37322E+18	İstanbul sözleşmesinin uygulandığı ülkelerde kadın cinayetlerinin olmadığını falan mı düşünüyorsunuz? Kadını istanbul değil islam sözleşmesi yaşatır. #istanbulsozlesmesi #Morardınızmı *"Do you think that there are no murders of women in the countries where the Istanbul Convention is implemented? The Islamic contract keeps the woman alive, not Istanbul. #istanbulsozlesmesi #Are you bruised"*	negative

1,37313E+18	Kadına şiddete tabiki karşıyım.Lakin istanbul sözleşmesi içine LGBT gizlenmiş sözleşmedir. LGBT kurana aykırıdır. Kesinlikle taviz verilemez. #İstanbulSözleşmesi *"Of course, I am against violence against women. However, LGBT is a hidden contract within the Istanbul contract. LGBT is against the Quran. Absolutely no concessions can be made. #IstanbulContract"*	negative
1,37307E+18	Varsayalım eve zamansız geldiniz ve karınızı biriyle yatakta yakaladınız ve bu durum karşısında dava açtınız. Mahkemede kadının sadece beyanı yeterli olurken sizin herşeyi delillerle kanıtlamanız laZımdı. Bu yüzden #istanbulsözleşmesi nin iptali en hayırlısı oldu.	negative
	#istanbulsoezlesmesi Veda hutbesinde tüm müslümanlara hitaben... Kadınlar size Allahın emanetidir sözü ile kanunlaştı, @RTErdogan Ömrün bereketli olsun İslam Halifesi	negative

Figure 1.Sentiment of Tweets Related to Istanbul Convention



Testing Models to Find the Best Fit

a. Feature Extraction Techniques

In the first part of the analysis, the text data was converted into a quantitative format using both the Count Vectorizer (Bag-of-words) and Tf-idf (Term Frequency-Inverse Document Frequency) with the scikit-learn library in Python to implement machine learning models.

b. Supervised Machine Learning Algorithms

To be able to find the best-fitted model, five commonly used supervised machine learning algorithms were tested in the annotated tweets data set. Since the whole data set is missing true labels, annotated data set was used for this part of the analysis. The data was divided into two as training (%80), and test data (%20) sets. Each model's performance was tested according to its metrics, which are precision, recall, accuracy, and f1 score. Five models were used for comparison, which are Multinomial Naive Bayes (MNB), K-nearest Neighbors (KNN), Random Forest (RF), Logistic Regression (Log.Reg.), and Support Vector Machine (SVC).

c. **Optimization with GridSearchCV**

After applying both feature extraction approaches (BOW, Tf-idf), hyper-parameter tunning implemented on each one separately. Default evaluation metrics were optimized using GridsearchCV for each model with each feature extraction techniques. This method is efficient to find the optimal combination of hyper-parameters.

d. **Evaluation**

For the evaluation part, f1 score was used as an evaluation criteria, since f1 score is better than other metrics when the data distribution is imbalanced. The distribution of positive and negative labeled tweets are quite different. That's why, f1 scores were compared with each other with each feature extraction technique and before and after optimization. Table 2 and 4 represents the values before hyper parameter tunning and Table 3 and 5 illustrates metrics after optimization.

Table 2. Evaluation Metrics for Different Supervised Machine Learning Models (BOW)

	Precision					Recall					f1- Score				
	MNB	KNN	RF	LogReg	SVC	MNB	KNN	RF	LogReg	SVC	MNB	KNN	RF	LogReg	SVC
Positive	0.69	0.83	0.71	0.76	0.63	0.91	0.36	0.83	0.84	0.94	0.79	0.50	0.76	0.79	0.76
Neutral	0.72	0.35	0.55	0.62	0.69	0.40	0.91	0.57	0.59	0.29	0.52	0.50	0.56	0.60	0.41
Negative	0.50	0.29	0.61	0.56	0.70	0.32	0.04	0.18	0.40	0.21	0.39	0.08	0.28	0.47	0.32
Accuracy											0.68	0.46	0.66	0.70	0.64

MNB: Multinomial Naive Bayes, **KNN:** k-nearest neighbours, **RF:** Random Forest, **Log.Reg.:** Logistic Regression, **SVC:** Support Vector Machine

Table 2 illustrates the evaluation metrics for five selected supervised machine learning models in which Bag-of-Words used as a feature extraction technique. The fact that bag-of-words (BOW) is a naive approach and it loses the semantic meaning of the words and tf-idf is much more mathematically elegant and more efficiently ignores common words, the results are better for tf-idf metrics (Table 4).

Table 3. Evaluation Metrics for Different Supervised Machine Learning Models after Optimization with GridSearchCV (BOW)

	Precision					Recall					f1- Score				
	MNB	KNN	RF	LogReg	SVC	MNB	KNN	RF	LogReg	SVC	MNB	KNN	RF	LogReg	SVC
Positive	0.69	0.62	0.61	0.76	0.70	0.91	0.75	0.97	0.84	0.85	0.79	0.68	0.75	0.79	0.77
Neutral	0.72	0.43	0.73	0.62	0.62	0.40	0.46	0.25	0.59	0.54	0.52	0.44	0.37	0.60	0.58
Negative	0.50	0.62	1.00	0.56	0.55	0.32	0.05	0.07	0.40	0.24	0.39	0.10	0.12	0.47	0.33
Accuracy											0.68	0.56	0.63	0.70	0.67

MNB: Multinomial Naive Bayes, **KNN:** k-nearest neighbours, **RF:** Random Forest, **Log.Reg.:** Logistic Regression, **SVC:** Support Vector Machine

Table 4. Evaluation Metrics For Different Supervised Machine Learning Models (TF-IDF)

	Precision					Recall					f1- Score				
	MNB	KNN	RF	LogReg	SVC	MNB	KNN	RF	LogReg	SVC	MNB	KNN	RF	LogReg	SVC
Positive	0.61	0.79	0.64	0.69	0.65	0.98	0.74	0.93	0.91	0.96	0.75	0.76	0.76	0.78	0.78
Neutral	0.77	0.53	0.71	0.69	0.78	0.25	0.59	0.36	0.46	0.37	0.37	0.56	0.48	0.55	0.50
Negative	0.00	0.35	0.70	0.59	0.70	0.00	0.37	0.15	0.24	0.15	0.00	0.36	0.25	0.34	0.25
Accuracy											0.62	0.64	0.65	0.68	0.67

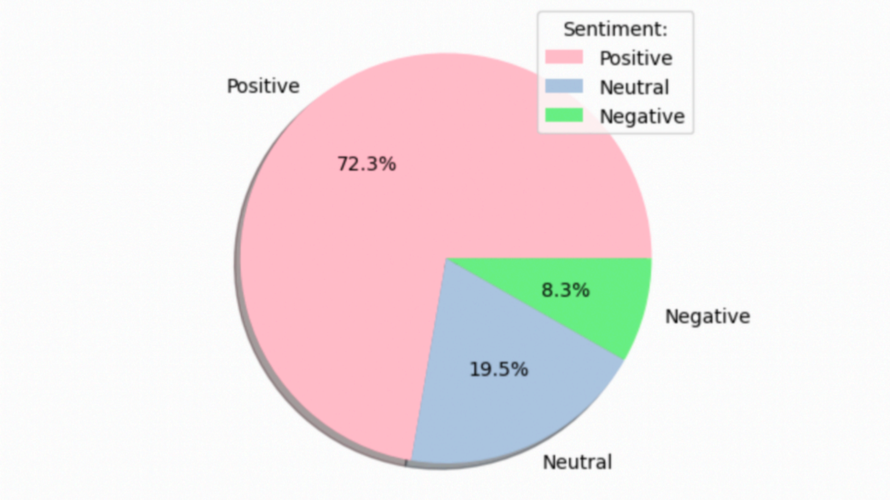
MNB: Multinomial Naive Bayes, **KNN:** k-nearest neighbours, **RF:** Random Forest, **Log.Reg.:** Logistic Regression, **SVC:** Support Vector Machine

Table 5 and Table 3 represent the metrics after optimization with GridSearchCV. When the default evaluation metrics were compared with optimized metrics, optimized versions gave better results. The tables were examined for each of the vectorization approaches after optimization by looking at their f1 score and Support Vector Machine with tf-idf as a feature extraction technique turned out to be the one with the highest f1 score. That's why, Support Vector Machine (tf-idf) has been selected as the best fit model for this analysis after optimization.

Prediction of Sentiment for Istanbul Convention Related Tweets Data Set

The Support Vector Machine model has been trained by using annotated tweets data set. The annotated tweets have been removed from the whole data set for this part. After getting predictions for the entire tweets, the new data set with predictions was imported. As shown in Figure 2, 72.3% of all tweets has positive sentiments, whereas 8.3% has negative sentiments. The tweets with neutral sentiment constitute the 19.5% of all tweets.

Figure 2.Sentiment of Tweets Related to Istanbul Convention After Training



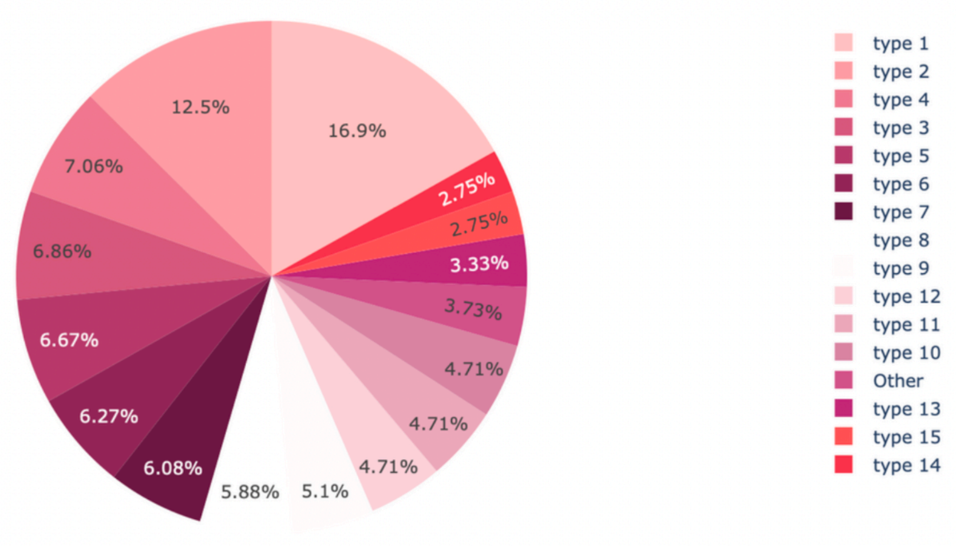
Topic Modeling for Each Tweets with Negative and Positive Sentiment

GSDMM mainly aims to detect topics in smaller documents than LDA, so it is more suitable for topic modeling of micro-blogging content such as tweets. That's why, GSDMM has been selected for the topic modeling. GSDMM generated a list of a total number of clusters with the top 20 most-recurring words and the number of times each of those words occurs in the topic.

GSDMM for Tweets with Negative Sentiment

Figure 3 represents clusters for topics created by GSDMM, which cluster has more tweets. It illustrates the importance of topics in the negative sentiment tweet data set. The types are made based on their importance and popularity. So, type 1 represents the most popular and important tweets, and the cluster with the most tweets is the most important (Figure 3.).

Figure 3. GSDMM for Tweets with Negative Sentiment



Some of the clusters selected from the data set which seem to indicate a topic compared to others and ranked based on their importance (Table 6.). Cluster 10 and 11 (which refer type 1 and type 2 in the pie chart) are the most important topics in negative labeled tweets. The fact that the words in those clusters are related to women, women's rights and the convention, this is not a surprising outcome. As expected before the analysis, Turkish family structure and Islamism also appear as a quite distinct category. The topics assigned to clusters based on the words that they include and occurrence of these words can be seen in the table below.

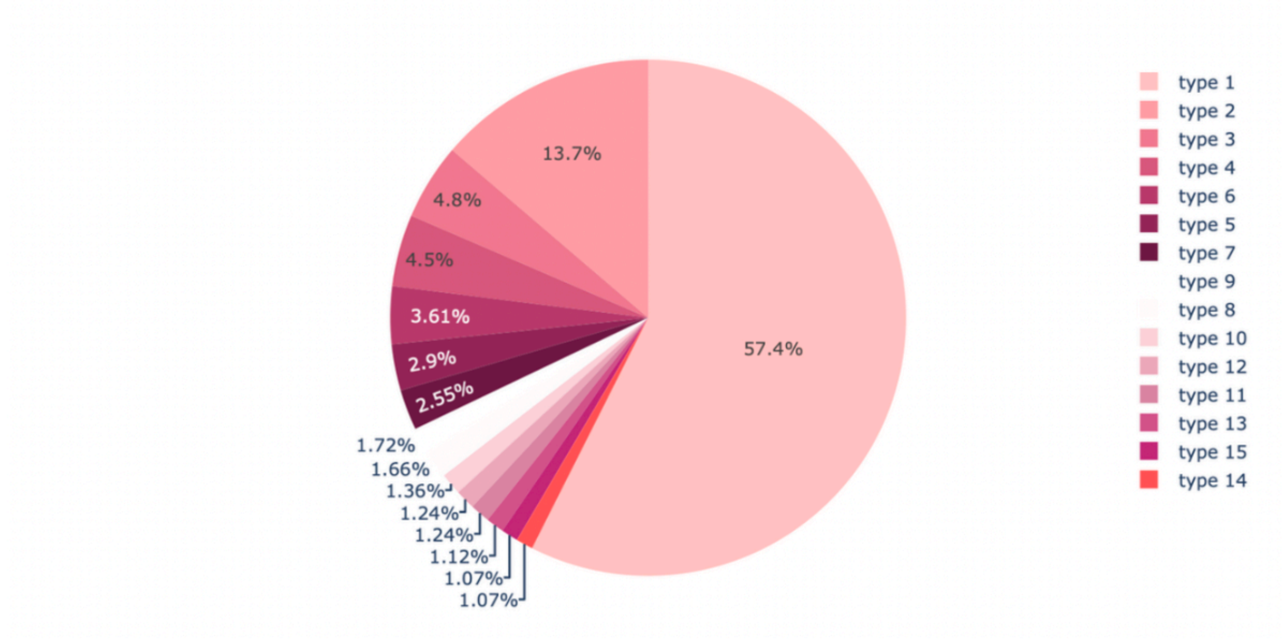
Table 6. Selected Topic Names and Words with Frequencies for Negative Labeled Tweets

Cluster No & Topic Name	Words & Frequencies
Cluster 10: Istanbul Contract	[('sözleşme', 80), ('stanbul', 58), ('bir', 33), ('olmak', 33), ('kadın', 29), ('hal', 12), ('demek', 12), ('biz', 11), ('desen', 11), ('var', 10), ('vermek', 10), ('iyi', 9), ('slâm', 9), ('kalmak', 9), ('avrupa', 8), ('devlet', 8), ('mi', 8), ('iş', 8), ('yok', 8), ('cinayet', 8)]
Cluster 11: Women and Rights	[('sözleşme', 61), ('kadın', 59), ('stanbul', 52), ('bir', 24), ('olmak', 24), ('var', 20), ('korumak', 19), ('tecavüz', 14), ('mi', 14), ('bu', 13), ('şiddet', 13), ('ülke', 12), ('demek', 12), ('cinayet', 11), ('taciz', 11), ('yok', 9), ('değmek', 9), ('desen', 9), ('kadar', 9), ('hak', 8)]
Cluster 13: Political and Cultural Aspects of Istanbul Convention	Cluster 13 : [('sözleşme', 19), ('stanbul', 13), ('olmak', 11), ('bir', 10), ('aile', 10), ('yapı', 8), ('çöp', 7), ('fatih', 7), ('ahlak', 7), ('millet', 6), ('var', 6), ('erbakan', 6), ('türk', 6), ('artık', 6), ('şükür', 6), ('önce', 5), ('lider', 5), ('atmak', 4), ('ayasofya', 4), ('şimdi', 4)]
Cluster 9: Islam and Turkish Family Structure	Cluster 9 : [('bir', 16), ('sözleşme', 12), ('aile', 11), ('olmak', 10), ('stanbul', 8), ('erkek', 8), ('elhamdülillah', 8), ('kadın', 5), ('ede', 5), ('müslüman', 5), ('feshetmek', 5), ('nesil', 4), ('ifade', 4), ('bitmek', 4), ('ge', 4), ('kadar', 3), ('yıkanmak', 3), ('söz', 3), ('türk', 3), ('belâ', 3)]

GSDMM for Tweets with Positive Sentiment

In the results of the GSDMM for positive sentiment tweets, type 1 represents the most popular and important tweets, and also the cluster that has the most tweets, the most important cluster for positive sentiment tweets as well (Figure 4.).

Figure 4. GSDMM for Tweets with Positive Sentiment



Some clusters from all clusters of positive labeled tweets that indicate a topic named and ranked in this part as well (Table 7). For positively labeled tweets, three topics were created based on the distinctiveness of the words in the clusters. The first topic represents the words that are associated with being a women in Turkey. In line with the literature, the second ranked topic in the table below represents one of the saying of Mustafa Kemal Atatürk, founder of Turkish Republic, which is “Kadınlarını geride bırakan toplum geride kalmaya mahkumdur.” (The society that leaves its women behind is doomed to be left behind!”). The last cluster that appeared as a distinct topic is related to oppression of women which is quite compatible with the issues in the Istanbul Convention.

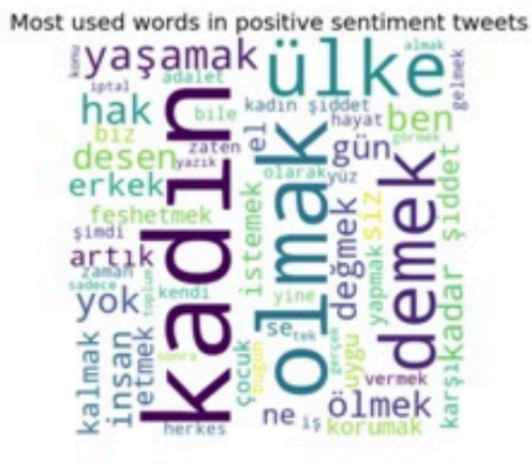
Table 7. Selected Topic Names and Words with Frequencies for Positive Labeled Tweets

Cluster No & Topic Name	Words & Frequencies
Cluster 4: Being a Woman in Turkey	[('sözleşme', 22), ('stanbul', 19), ('bir', 12), ('zulüm', 8), ('demek', 6), ('kadın', 6), ('kanu', 6), ('olmak', 6), ('kalmak', 6), ('söylemek', 6), ('insan', 5), ('çekmek', 5), ('yaşamak', 5), ('kendi', 5), ('akp', 4), ('bu', 4), ('önce', 4), ('zaten', 4), ('utanmak', 4), ('biz', 4)]
Cluster 11: Saying of M. Kemal Atatürk	[('kadın', 51), ('geri', 36), ('sözleşme', 20), ('bırakmak', 18), ('kalmak', 18), ('stanbul', 16), ('toplum', 15), ('kemal', 15), ('cinayet', 14), ('yaşamak', 13), ('politik', 11), ('atatürk', 10), ('iş', 9), ('şiddet', 8), ('karşı', 8), ('mahkum', 7), ('mustafa', 7), ('not', 7), ('safe', 7), ('for', 7)]
Cluster 2: Oppression of Women	[('sözleşme', 22), ('stanbul', 19), ('bir', 12), ('zulüm', 8), ('demek', 6), ('kadın', 6), ('kanu', 6), ('olmak', 6), ('kalmak', 6), ('söylemek', 6), ('insan', 5), ('çekmek', 5), ('yaşamak', 5), ('kendi', 5), ('akp', 4), ('bu', 4), ('önce', 4), ('zaten', 4), ('utanmak', 4), ('biz', 4)]

Results & Discussion

Word Clouds for the topics in positive sentiment tweets

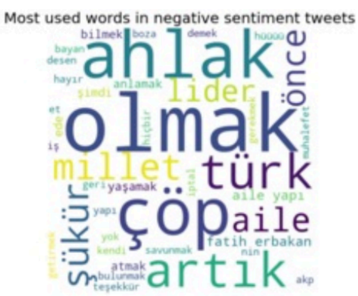
After applying topic modeling, word clouds are generated for each labeled tweet. For positive sentiment tweets, it can be said that one word cloud is strongly associated with one side of the debate in which people want convention back since their ideology taking its roots from Kemalist populism. In this word cloud, there are words such as “Kemal”, “geri bırakmak” (recede), “geri kalmak” (left behind), “Kemal Atatürk”, “mahkum” (doomed), “Mustafa Kemal”. It is strongly related to the literature since most people who support the Istanbul convention are seen as Kemalists. Those people are generally using one of the most popular statements of M. Kemal Atatürk, founder of the Turkish Republic, which is "Kadınları geride bırakan toplum, geride kalmaya mahkumdur" ("A society that leaves women behind is doomed to be left behind."). The fact that those words were appeared in one of the word clouds confirmed the findings in the literature. Since those people see the ruling party as responsible for the increasing femicide in Turkey and the abolishment of convention; words like "Katil" (murderer), "cinayet" (murder), akp (Justice and Development Party), "ahmak" (fool), "cinayet politiktir" (murder is political), and so on are seen. In addition to that, to be able to show their reaction to the withdrawal of the convention, words like "feshetmek" (abolish), "utanmak" (be embarrassed), "iptal" (cancel), "susturmak" (silence), "geri çekmek" (drawback), "yuh" (boo!), "ölmek" (die), "adalet" (justice), "şiddet" (violence), "zulüm" (oppression), "terk etmek" (abandon) vs. used frequently in the tweets that support Istanbul convention.



Word Clouds for the topics in negative sentiment tweets

Negative opinions concerning the Istanbul convention are mainly related to the disintegration of family unity, violation of Islamic rules, and Turkish culture. In light of the debate, topics after topic modeling have resulted into words related to Islam. Those words such as "müslüman" (Muslim), "islam" (Islam), "aile" (family), "ahlak" (morals), and "aile yapısı" (family structure), can be seen in the word clouds.

In addition, people who supported the withdrawal of the Istanbul Convention wanted to show their thanks and gratitude to the ruling party. “Amin” (amen), “reis” (nickname for president), “ooohhh”, “şükür” (glorification), “elhamdülillah” (alleluia!), “teşekkür” (thanks), “Fatih Erbakan” “cumhurbaşkanı” (president), “cb”(abb. for president) are the words that show this gratitude as they can be seen in the word clouds below.



Since the convention is associated with the LGBTGIA+ community from people who are against Istanbul Convention and the fact that many scholars and media figures claim that it is not suitable for Turkish culture; the convention is seen as trash, pervert, and deviant. Therefore, the words such as "çöp" (garbage), "eşcinsel" (homosexual), and "lgbt" in the word clouds represent negative opinions towards the convention.

Social and political debates are reflected in social media with the rise of social media usage. The digital sites shapes the everyday discourses and the public atmosphere and also political structure the content in online platforms. The reflection of this debate has appeared on social media, particularly on Twitter. This study has proved that social media is an essential tool for understanding political and social changes in society and how people react to those changes.

Social media is a place in which there is the freedom of speech and expression, however, it is also a place for public debate polarization, aggressive communication, and online harassment. The opposition towards feminist activism in the digital sphere causes the spread of anti-gender ideologies, and also takes part in the normalization of radical views. In the study, the topics for the negative and positive sentiment tweets clearly showed the topics debated. While tweets with negative sentiment are strongly associated with the conservative opposition of ruling party and the ideas of conservative politicians, tweets with positive sentiment are primarily based on the Kemalist populist perspective. The fact that the study's result is in line with the literature shows that social media is the successful representation of what happens in the society. This study underlines the importance of using social media to be able to understand the tendencies in the society better and paves the way for future studies.

Pitfalls and Future Suggestions

This study can be improved and redone by taking granted the pitfalls and challenges. The challenges can be seen in many parts of this study. First of all, Twitter data was used for the analysis. The nature of tweets creates difficulty in meaning extraction since they include short text, and lack the information regarding visuals, hyperlinks, or quoted tweets. In addition to that Twitter data is quite noisy which causes poorly performed models. The other limitations was in the annotation part of the analysis. The fact that it is hard to understand the context of within which a tweet has been sent makes the sentiment extraction difficult. In addition to that users' background information is missing, therefore, it is hard to detect irony within the tweet. Besides, only one annotator annotates these tweets, so, this might cause some errors and biases in the data set. More than one annotator can be used to strengthen the analysis results. The labels can be justified by pair wise comparisons. Another issue is that the data set is limited due to time concerns, so limited data set might restrict the diversity of topics in the topic modeling part. A more extensive annotated data set will be more efficient for the model to predict the whole data set and the diversity of the topics in the analysis. Most of the challenges of this study were in the preprocessing part. There are not enough well-developed preprocessing methods and libraries for the Turkish language, and the library used in this analysis, nltk, failed to cover all stop words in the Turkish language. Besides, due to the Turkish language problem, the failure in normalization and lemmatization, there are lost of meaningless words. Another problem with Turkish language is that it prevents the usage of globally accepted packages such as VADER and Textblob for sentiment analysis. Furthermore, NLP starts with a morphological process, and there are many challenges in Turkish due to its complex morphology.

This study is essential for understanding the political and social changes in society and people's tendencies and reactions to those changes. The results of this study paves the way for future studies.

The fact that social media gives insights about the society makes it crucial to use in societal studies.

This study will also helpful to make policy changes which feels the pulse of society.