Tensions in Artsakh:
Sentiment and Topic Modelling of Tweets Surrounding Conflict in Nagorno-Karabakh

Ani A. Parnagian

IPHS 290 Cultural Analytics (Fall 2022), Prof Elkins and Chun, Kenyon College

Abstract

What is the general sentiment and key words brought to light by Twitter discussions surrounding the war and tensions in Nagorno-Karabakh? What are sentiments surrounding the Prime Minister of Armenia and President of Azerbaijan, Nikol Pashinyan and Ilham Aliyev respectively? This project hope to understand and publicize the online discussions surrounding the wars, and how they have changed from tensions in 2016 with the Four-Day War compared to the War in Fall 2020 and the change of power in the region.

Introduction

In 1923, the Soviet Union established the Nagorno-Karabakh Autonomous Oblast, an area home to 95 percent ethnically Armenian population, but in the Azerbaijan Soviet Socialist Republic. As the Soviet Union began to collapse in the late 1980s and early 1990s, tensions in the area rose as Nagorno-Karabakh (also known as Artsakh) declared its intention to join the Republic of Armenia in 1988, and then declared their own independence in 1991. War erupted at the time, resulting in 30,000 casualties. By 1992, Armenia had gained control of Nagorno-Karabakh. In 1994, Russia brokered a ceasefire. There was general peace in the region for two decades, will intermittent clashes and attacks in the Area. A Four-Day War broke out in the region in April 2016, but a full-blown war broke out in September 2020. A deal was brokered in November 2020, with Azerbaijan taking over much most of the territory. Tensions are as high as ever, and there are continuous tensions. In just the last few days, Azerbaijan has closed the only road connecting Nagorno-Karabakh to Armenia, as well as cutting the natural gas supply to Nagorno-Karabakh, leaving 120,000 people without the ability to heat their homes or leave the region.

Methodology

For this project I utilized sentiment analysis of tweets to receive my corpus of tweets. The keywords I used were "Artsakh War", "Nagorno Karabakh," "Pashinyan," and "Aliyev."

I installed pandas, snscrape, as well as cleaning tools for the texts that cleaned emoticons as well as contractions to help clean the tweets as best as the program can. I ran the sentiment I then placed my csv file into code for topic modelling. First, this notebook cleans the text file, removing stopwords (eg. 'and,' 'or,' 'what') and then lemmatizes words: this means finding the stems of words so that there is not unnecessary repetition (eg. run and running are considered the same for the sake of the topic modeling. This notebook then uses Gensim Latent Dirichlet Allocation Model. LSA is a way of discovering topics within text. It is able to search through the corpus of tweets and places words found in the tweets in different categories.
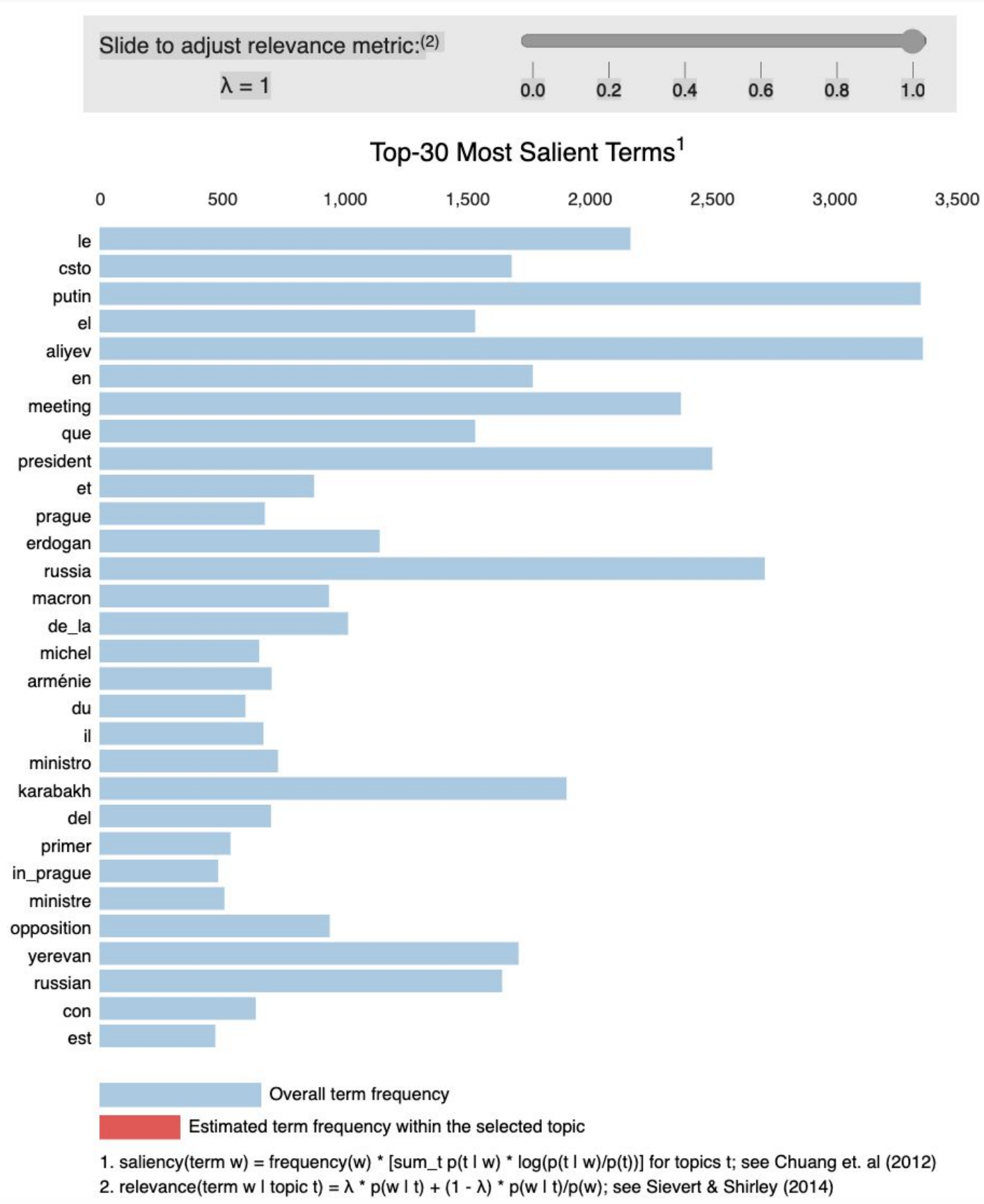
Results

My first search term "Artsakh War" yielded results stretching back to January 2021, with the greatest amounts of tweets generated then, just as the war was ending. Finally, my searches on "Pashinyan" and "Aliyev" aimed to see sentiment surrounding leadership during and after the war.

A lot of this running was in foreign languages, making the topic modelling a bit more confusing to understand. Almost all of the tweets surrounding Aliyev were in Azerbaijani or Turkish, which demonstrates that many of the people having conversations about Aliyev are likely in support of him.

In terms of the discussions surrounding Nikol Pashinyan, mention Putin and Erdogan, as well as Aliyev, highlighting the issues he had navigating the international politics of the region. The words "meeting" and "opposition" are also mentioned highlighting the anger many Armenians had with his handling of the situation.
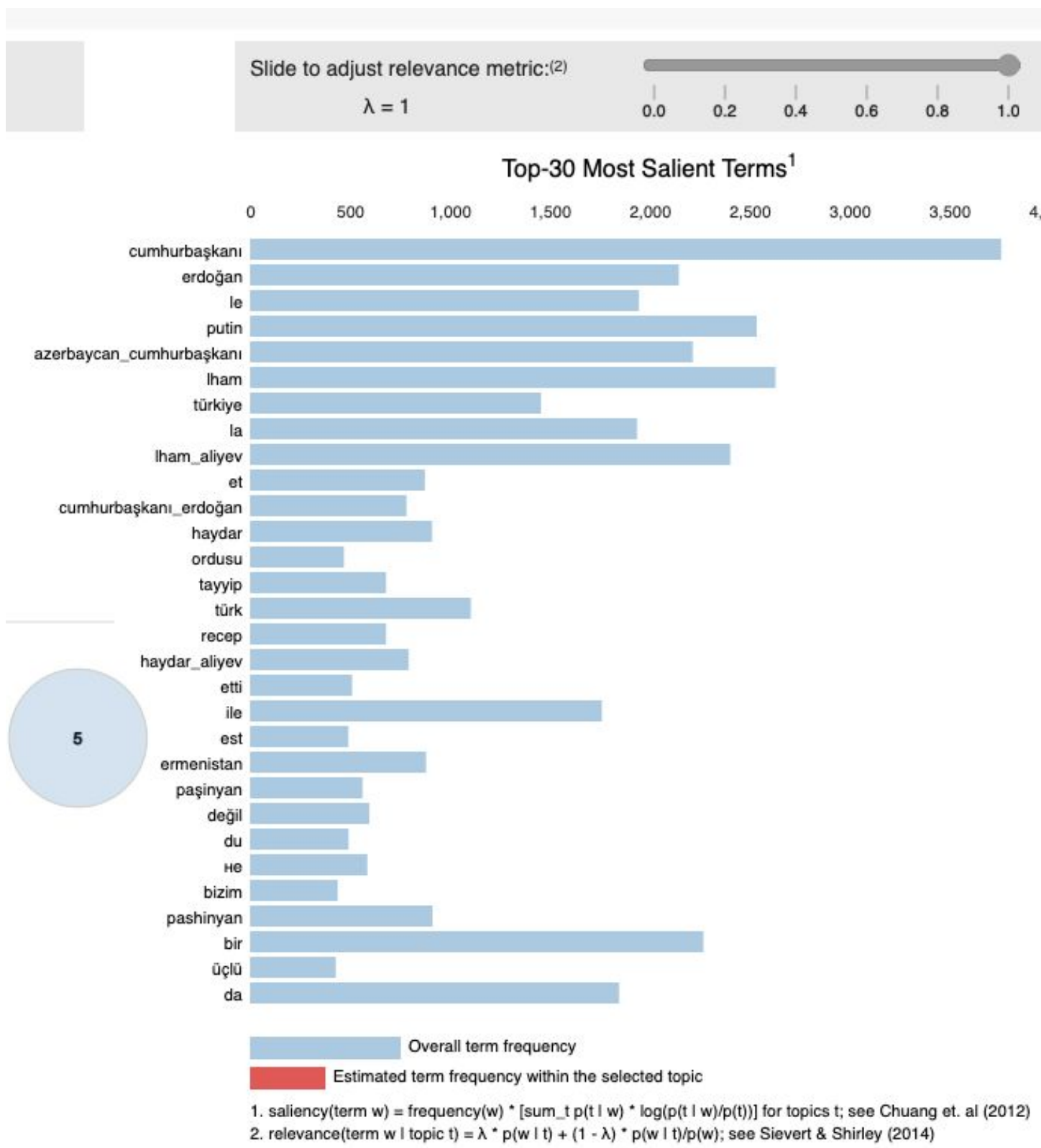
Conclusion

The sentiment analysis of this projects, as seen below using RoBERTa, demonstrates incredibly negative sentiment surrounding this issue. This makes sense given the violence in the area. Sentiments this negative prove to me that many people understand the injustice happening in the area while new outlets refuse to give the correct information. While there are no huge surprises in this research I think it demonstrates how forgotten the atrocities happening in the area have been largely forgotten.
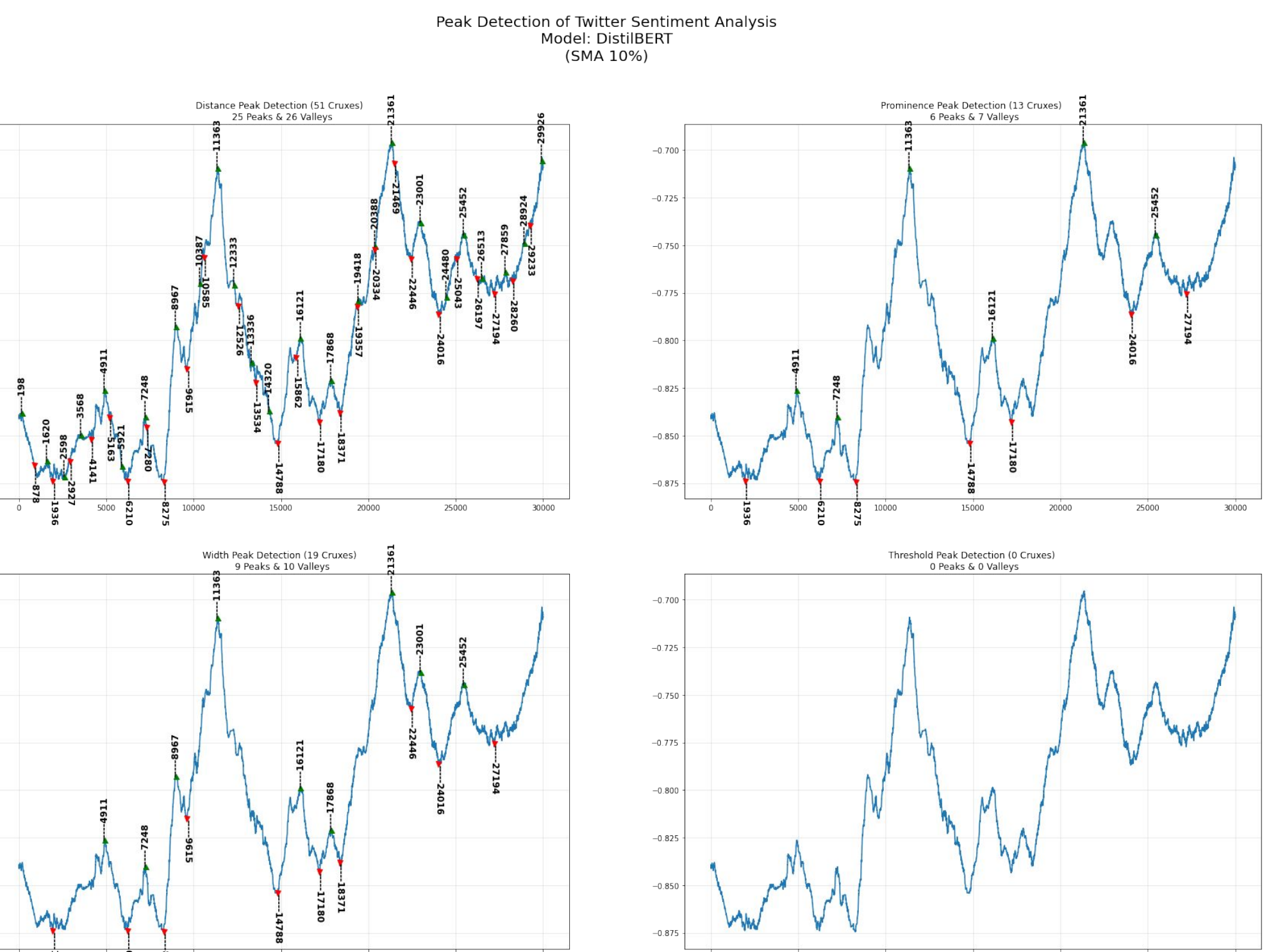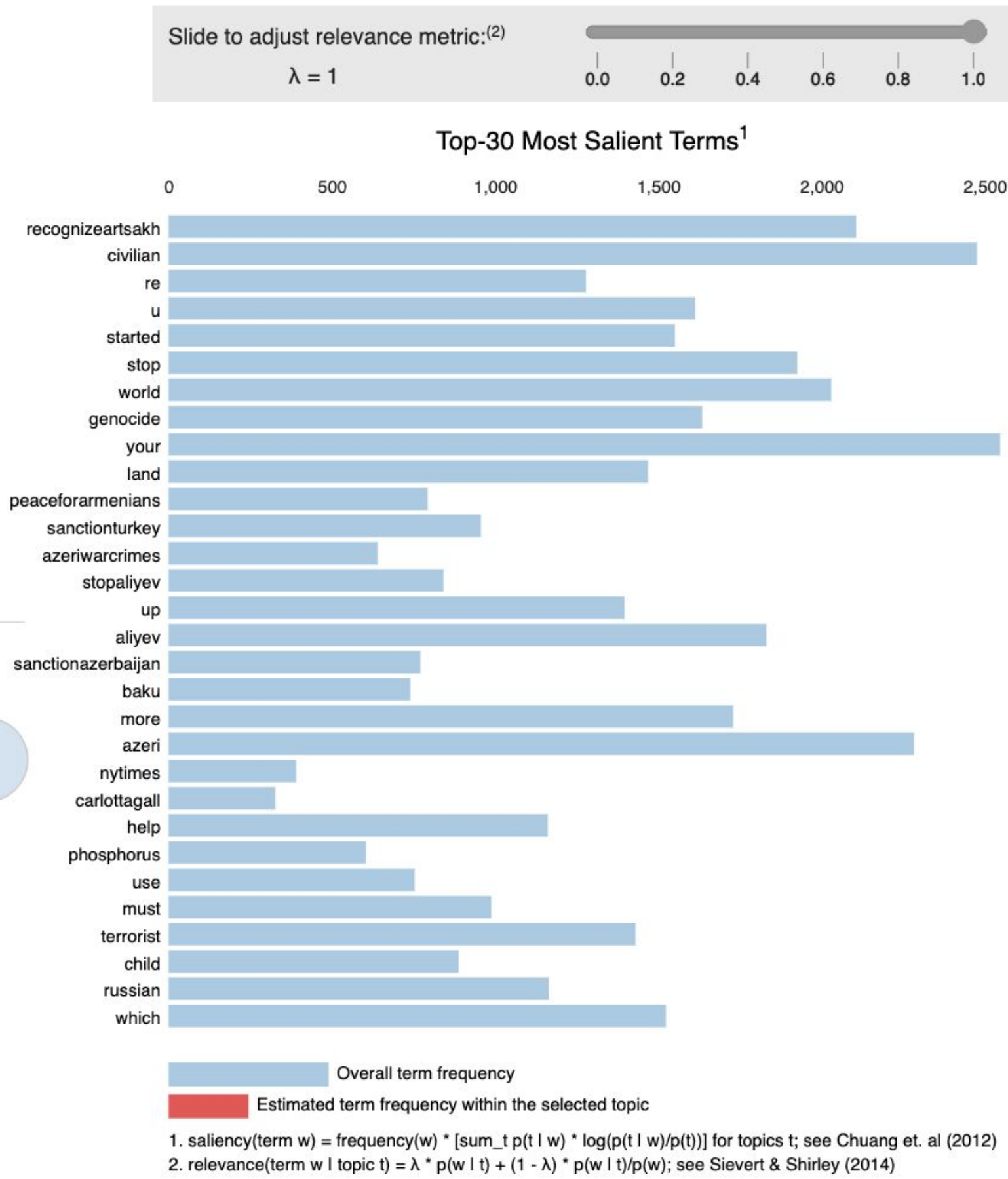

Peak Detection of Twitter Sentiment Analysis
Model: DistilBERT
(SMA 10%)

References

Chorbajian, Levon, Patrick Donabédian, and Claude Mutafian. *The Caucasian Knot: The History & Geopolitics of Nagorno-Karabakh.*

"Introduction to Latent Dirichlet Allocation." n.d. Blog.echen.me. Accessed December 14, 2022. https://blog.echen.me/2011/08/22/introduction-to-latent-dirichlet-allocation/.

Notebooks by Professor Jon Chun: Sentiment Analysis of Tweets, and Topic Modelling for Tweets

Acknowledgements

Nikol Pashinyan Topic Modelling

Ilham Aliyev Topic Modelling

Artsakh War Topic Modelling