# The US 2020 elections through the eyes of TikTok

Ravinithesh Annapureddy ravinithesh.annapureddy@epfl.ch EPFL Lausanne, Switzerland Maxime Jan maxime.jan@epfl.ch EPFL Lausanne, Switzerland Elisa Michelet elisa.michelet@epfl.ch EPFL Lausanne, Switzerland



### Figure 1: Joe Biden (left) and Donald Trump (right)

# ABSTRACT

TikTok is a video-based social media application that has recently garnered popularity and saw a steep increase in users. It was the most downloaded app in 2020, a year that has seen some significant times in the history of the US. One of them is the presidential election, and this study aims to comprehend the political communication about the election on TikTok. A set of videos that contained the hashtags from a predetermined set were collected and used in the analysis. The hashtags in these videos were investigated using statistical tools, network analysis and topic modelling. The results suggest that the communication is interactive, and the young people are talking about issues rather than parties or personalities. However, they season their message with fun and trendy challenges. Additionally, two categories of hashtags emerged. The first set contained the seasonal ones related to the election, and the other had hashtags with a theme attached to them.

## **CCS CONCEPTS**

Human-centered computing → Social network analysis; Collaborative and social computing theory, concepts and paradigms.

# **KEYWORDS**

TikTok, hashtags, political communication, topic modelling, network analysis, social media, US politics, US election

CSM'21, Spring, 2021, EPFL

© 2021 Copyright held by the owner/author(s).

### **ACM Reference Format:**

Ravinithesh Annapureddy, Maxime Jan, and Elisa Michelet. 2021. The US 2020 elections through the eyes of TikTok. In *CSM '21: Computational Social Media course, Spring 2021, EPFL, Switzerland.*, 9 pages.

## **1** INTRODUCTION

Campaigning is at the heart of an election process, and delivering the message to the voters is the most crucial component of it. Posters, flyers, ads in newspaper, television and the web were important tools to send this message. However, the 2008 Unites States of America (US) presidential election has seen the use of social media for the first time when Barack Obama widely used Twitter and advertised through YouTube videos [10]. The advantage of social media is that it not only allows to deliver the messages but to deliver them in a targeted fashion. This focused communication could influence the voter more in comparison to a global message. Thus social media became more significant as the years passed in various campaigns including politics in and outside of the united states [11].

The United States of America is the world's oldest democracy, the largest economy and its president is often referred to as the leader of the free world. These reasons, coupled with other geopolitical factors, explains why the US presidential election, which is held once in four years, receive global attention and is also relevant to the international population [8]. During the term of Donald Trump, there was a significant shift in the political discourse of the United States and the election at the end of his term has created much-needed hype. Along with it came many other era-changing moments in US history such as the Black Lives Matter protests and the COVID-19 pandemic. All these events made the 2020 US presidential election, between the prime contenders Joe Biden and Donald Trump from the Democratic and Republican parties respectively, one of the most anticipated ones in history. Given the rise of social media usage and the importance of the 2020 presidential

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

### CSM'21, Spring, 2021, EPFL

election, this work tries to understand the US election through the latest social media mania: TikTok.

TikTok is a mobile app that is used for making and sharing short videos. It was initially launched as *Douyin* in September 2016, in China. The following year, in 2017, the app was launched by *ByteDance* for markets outside of China [14]. As of January 2021, there are more than 2 billion downloads [4] and 689 million monthly active users outside China [9], making it the 7<sup>th</sup> largest social media platform [6].

If Twitter made a shift from blogging to micro blogging, one can say TikTok brought the shift from videos to micro-videos. Users on TikTok can make short videos and post them on the site to share them with others. Another analogy that can be drawn with Twitter is about the size/length. Initially, the maximum allowed length of a tweet was 140 characters and it is now 280 characters. Similarly in TikTok, the initial video limit was 15 seconds and it is now extended to 60 seconds.

The growth of TikTok is rapid [6], especially, the number of monthly users of TikTok has increased steadily in the first half of 2020 [9]. Particularly in the US, the number of monthly active users grew to more than *91 million* by June 2020 compared to the last know number of *39 million* in October 2019 [9]. It was the time when many countries in the world have imposed lockdowns and confinements in response to rapidly growing COVID 19 cases [1]. The lockdowns, homeschooling and work from home options might have enabled people to use these kinds of apps to spend their time. At the same time, the user composition of TikTok in the US shows interesting patterns. While 25% of the users are teens whose age is less than 19, around 45% of them are between the ages of 20 and 40 [12], who are not too old. Along with revealing that TikTok is very popular among younger people, these statistics also indicate that TikTok is going to be the opinion platform for the new generation.

It is generally thought that the younger people are not keen on politics. However, multiple factors have made the 2020 election even more interesting than they usually are. Starting from the actions of the previous president, the management of COVID, the economy, wide anti-racism movements, trade wars, and people staying at home has made the elections a common talking point. As young people grew more interested in the election and TikTok was one of the major platforms that had the presence of these users, the app contributed its share in shaping the presidential election in 2020. Thus, we shall try to answer the following questions through this project:

**RQ 1:** How did the momentum to the election culminate on TikTok over time?

**RQ 2:** What kind of hashtags are used together and how were they used in the videos?

**RQ 3:** What are the topics of discussion in the election discourse through TikTok?

To answer these questions, first, the metadata of the TikTok videos related to the 2020 election was collected following European Union (EU) and US laws. Then, the method of hashtag appearance count in videos was used to answer RQ 1. Hashtag co-occurrences and hashtag network analysis for hashtag community detection were used to answer RQ 2. Lastly, videos graph partitioning for topic modelling was used to answer RQ 3.

The rest of the paper is organised as follows. The next section describes the data collection and ethical concerns related to it. Section 4, 5 and 6 describe the individual methods and results obtained as answers to each research question. Conclusions and limitations are drawn in the last sections.

# 2 DATA

This work uses the metadata of TikTok videos containing specific hashtags related to the 2020 US presidential elections. In particular, all the hashtags of each of these videos were queried using an API in accordance with the laws on data regulation.

# 2.1 Ethic and Juridic Concerns on Data Collection

The collection of data had to be done in compliance with the national and international laws on data privacy. As the research project was done at EPFL, in Switzerland, it had to follow Switzerland and EU's General Data Protection Regulation (GDPR) laws. However, because most of the subjects i.e., the creators of the videos used in this project, will be citizens of the United States, US laws must also be followed. Three issues with the data collection must be addressed: the information obligation, the use of automated scripts to collect data, and the creation of new sensitive data.

2.1.1 Information Obligation. The GDPR states in its texts an information obligation: "every participant in a study should be informed" (Art. 14 par. 1 GDPR). There are however two exceptions for this directive. The first one is if the users were already informed (Art. 14 par. 5 lit. a GDPR) that their data can be used by others. In the case of the project, TikTok users generally do not know up to which extent the information they share on the public platform will be used by others. So the first exception cannot be applied to this project. Another way to be exempted from the information obligation is when the task of informing the subjects would involve a disproportionate effort (Art. 14 par. 5 lit. b GDPR). In this study, the number of participants is indeed too large, and because the data is collected through an intermediary, that would mean having to look actively for the identity of each user. This would not only be too much effort but also go against the terms of use of TikTok. Thus, under this exception, the user's approval of the data usage was omitted.

2.1.2 Data collection API. The project used an unofficial API TikTokApi<sup>1</sup> for collecting the required data. The use of external libraries to collect TikTok data is not clearly authorised nor discouraged by the TikTok Developer Terms of Service. The US laws specifically forbid the usage of automated scripts for data collection *18 U.S. Code § 1030.* However, as the project was based in Switzerland the US law on using automated scripts was not applicable. What applied here were the Swiss laws, that authorise using automated scripts for data collection if it is not for commercial purposes. Hence, using a third-party interface for this project was then acceptable.

2.1.3 Sensitive data. The last sensitive part of this project was about the kind of data that will be collected, and what was planned to do with it. At EPFL, the processing of personal data in the context

<sup>&</sup>lt;sup>1</sup>https://pypi.org/project/TikTokApi/

of scientific research was authorised (*Art.36c ETH Act*), providing a solid ground to work on. Nevertheless, users of TikTok potentially are children or teenagers, which makes them vulnerable subjects. Not only are they vulnerable, but the project would also seek sensitive information: their political opinions. That is why particular care was taken at not studying the users individually, but rather looking at the overall use of hashtags, and the topics discussed. No new sensitive information - like the political orientation, or the belonging to a community - was inferred from the data.

# 2.2 Data Collection

To collect metadata from TikTok, a list of 16 relevant hashtags was first manually compiled (Appendix B). To balance the dataset, 6 hashtags associated with the democratic party, and 6 others related to the republican one were chosen. The last 4 hashtags were neutral as they concern the elections in general. The metadata of the 500 most popular videos containing each of these hashtags was then queried with API. This resulted in 8000 videos and after discarding the videos outside of 2020 and dropping the duplicates, only the videos' date, list of hashtags, and ID were kept. The hashtags fyp and xyzbca were ignored during the experiments as they were considered as stop words because used in almost every video. fyp is short for For You Page, the main page where the videos are shown to the users and the users believe that using *fyp* and *xyzbca* will trick the TikTok algorithm to show their videos on other users landing page. In total, the metadata of 5232 unique videos were collected. Table 1 shows a sample of 3 entries of this dataset.

hashed Video ID	hashtags	date
-49374569309301	[aussie, biden2020,	2020-11-07
	bidenharris2020, straya]	01:51:07
28378852204966	[biden, biden2020, elec- tion2020, fyp, fürdich]	2020-11-06 18:11:03
-76166906306051	[biden2020, fyp]	2020-10-19 18:41:40

Table 1: First three rows of the dataset of TikTok's metadata

# 3 MOMENTUM TO THE ELECTION ON TIKTOK

Before studying further questions, it is crucial to understand to what extend did the 2020 US presidential election get traction on the TikTok platform. Therefore, this first analysis investigates when and how the interest for this event evolved on this social media platform.

### 3.1 Methods

Three different methods combining the dates and hashtags are applied to visualise the evolution of hashtags' distribution over time. 3.1.1 Frequent hashtags' distribution estimation. As a means to describe the general trend of hashtag use related to the 2020 elections, their distribution is approximated over time with a Kernel Density Estimation (KDE). Such a method allows representing the data using a continuous probability density curve. First, each individual hashtag is associated with its date of publication and the ones appearing less than 600 times were discarded. Then, KDEs approximate the occurrences of each of these remaining hashtags per day. Here, the KD estimator  $\hat{f}(\mathbf{x})$  is defined as [7]:

$$\hat{f}(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\sqrt{2\pi}} \exp(-\frac{||\mathbf{x} - \mathbf{x}_i||}{2})$$
(1)

where  $x_1, x_2, ..., x_N$  represent the number of each remaining hashtag per day. The KDEs are then stacked to visualise the general trend.

*3.1.2 Peaks of use.* The number of videos per day is aggregated and then sorted decreasingly, which highlights the days with particular peaks of political TikToks. The top 10 peak days were then linked with actual events of the 2020 US presidential election to observe whether these peaks correspond to reactions to these events.

3.1.3 Frequent hashtags' relative popularity. To understand the relative usage of hashtags, their appearance count was computed and only those that appeared at least 600 times overall were considered. These counts were then aggregated per month and each hashtag's popularity was computed as its fraction of use relatively to all considered hashtags.

### 3.2 Results

3.2.1 *Distribution estimation.* The stacked KDE can be observed in Figure 2. The 7 frequent hashtags that were retained were *voteblue*, *trump2020, trump, election2020, election, biden2020, and biden.* It can be observed on this plot that the density steadily grows until the day of the election, reaching a maximum near the beginning of November. After this peak, this density rapidly decays until the end of the year. It must be noted that while the dataset does not contain any data of 2021, the KDEs extrapolated the distributions until March of that year.

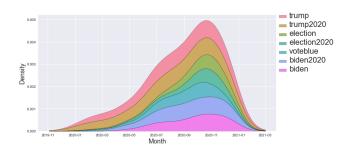
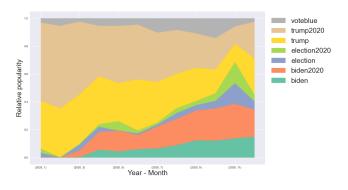


Figure 2: Stacked KDEs of the most frequent hashtags

*3.2.2 Peaks of use.* The top 10 peak days are given in Table 2. These dates can be linked with the following events.

CSM'21, Spring, 2021, EPFL



# Figure 3: Plot of the relative popularity of the hashtags appearing more than 600 times

- November 2-8: 7 out of the 10 peak days happened around the Election Day, in a short time frame slightly shifted towards the days where votes were being counted.
- September 30: This peak falls on the day of the infamous first presidential debate between Biden and Trump.
- October 27: This peak happens on the same day as the Senate controversially confirmed Amy Coney Barret for the Supreme Court of the United State only one week before the Election Day.
- October 23: This peak happens on the day of the second presidential debate.

Date Number of videos poste	
2020-11-04	395
2020-11-07	278
2020-11-05	260
2020-11-03	243
2020-11-06	201
2020-09-30	145
2020-11-08	127
2020-11-02	103
2020-10-27	58
2020-10-23	58

Table 2: Top 10 days with the most number of videos postedcontaining political hashtags

3.2.3 Frequent hashtags' relative popularity. The relative popularity of the frequent hashtags enumerated in section 3.2.1 is displayed in Figure 3. The popularity of the democratic hashtags is close to zero at the beginning of the year with Trump related ones having almost 100% of the popularity. Biden related hashtags only gain momentum in the month of June, but rapidly catch up with Trump related ones. At the time of the elections, democratic hashtags even have higher popularity than the republicans.

### 3.3 Discussion

From these three analysis, it can be affirmed that the TikTok platform showed a deep interest in the 2020 US presidential election. Although the interest in this event was rather low at the beginning of the year, it gained a lot of traction since June. Moreover, the above results showed that the platform remained updated with the ongoing events of the elections, strongly reacting to the important and controversial ones.

Finally, the last analysis revealed a trend where Trump's popularity was extremely high at the beginning of the year and only decreased, while Biden's popularity steadily grew to finally catch with Trump. However, this result must be nuanced by two factors. First, as the previous analysis revealed, the interest in the elections only started gaining traction in June. This means that while Trump had a high relative popularity, it was only on a small number of videos. Second, Biden's apparent unpopularity at the start can be explained by the fact that the democratic primary was still ongoing and he did not become their candidate before April. Therefore, this may indicate that Biden's campaign was not a subject of interest on TikTok at all.

# 4 CO-OCCURRENCES OF HASHTAGS

The second question delves into the link between hashtags. The goal is to determine whether some hashtags are strongly correlated by being used often together and what communities can be inferred from those.

## 4.1 Methods

4.1.1 Co-occurrence matrix. In this first analysis, the instances of pairs of hashtags appearing in the same videos are counted via a co-occurrence matrix. With N different hashtags in the dataset, this matrix has a shape NxN. Each row and each column can therefore be labelled by a distinct hashtag. Then, each element  $m_{i,j}$  is equal to the number of videos where the hashtag of row i appears along with the hashtag of the column j.

As a toy example, let's. consider a dataset with 3 videos, containing respectively the hashtags [A, B], [B, A] and [C, B]. The co-occurrence matrix resulting from this dataset would be:

	А	В	С
A B	$(^{0}$	2	0)
В	2	0	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$
С	10	1	0)

Such a matrix was computed for all the videos' hashtags as was then truncated in order to only keep the hashtags with more than 2000 co-occurrences.

4.1.2 Hashtags communities. The second analysis seeks to find communities in the network of hashtags. To do so, a network of hashtags was created where each node represents a hashtag, an edge between two nodes exist if the hashtags occur together in a video and the edge weights are the number of videos they appear together. Only hashtags that have appeared at least 50 times were used. Then, the Clauset-Newman-Moore greedy modularity minimisation method was applied to this network [5]. This algorithm starts with each node being its community, and greedily merge the pair of communities that increase the modularity until no such pair exist.

The US 2020 elections through the eyes of TikTok

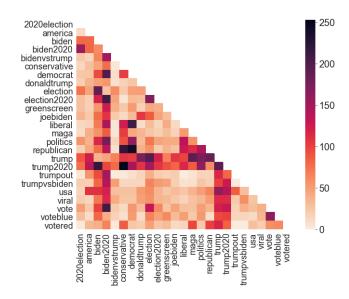


Figure 4: Bottom half-triangle of the truncated cooccurrence matrix

#### 4.2 Results

4.2.1 Co-occurrence matrix. The truncated co-occurrence matrix can be seen in the form of a heat map in Figure 4. As the matrix is symmetric, only the bottom half-triangle is displayed. On this plot, the darker an element, the more co-occurrences there are between the hashtags of the corresponding row and column.

4.2.2 Hashtags communities. The Clauset-Newman-Moore method detected two distinct communities, which are visualised in Figure 5. On this plot, the most frequent hashtags are grouped and coloured relative to the community they belong to.

#### 4.3 Discussion

4.3.1 Co-occurrence matrix. The first striking observation that can be made is that the Biden related hashtags often appear along with the Trump related ones. Indeed some of the darker spots in Figure 4 reveal a high co-occurrence of trump2020 and biden, as well as republican and democrat. Moreover, trumpout appears to be used along with both parties' hashtags. However, other hashtags, such as maga or votered seem to be more co-occurring with republican specific hashtags. Overall, this analysis reveals that it is difficult to observe clear groups of hashtags, as they are often mixed together. This result could indicate that TikTok's videos are more about the elections in general rather than used by a particular party for their campaign, or that the users choose many hashtags only to increase the reach of their videos.

4.3.2 Hashtags communities. The two communities of hashtags as seen in Figure 5 can be interpreted as follows: the hashtags coloured in blue as seasonal ones related to the election and those in red as those that are used stable across time with an associated theme. Looking deeper, the hashtags trump2020, biden2020, uselections, 2020elections, kamalaharris, bidenharris, trumpout relating to election and homeoffice referring to remote work during the COVID-19



CSM'21, Spring, 2021, EPFL

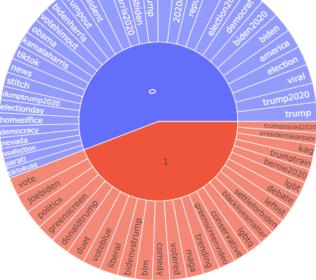


Figure 5: Hashtags communities from the Clauset-Newman-Moore algorithm

pandemic appeared in the blue part of the circle and these were popular during the 2020 elections. In the contrast, the other set contains hashtags such as maga, kag, liberal, blm, lgbtq along with common TikTok hashtags such as duet, trending, greenscreen. These hashtags have a theme or persona attached to them.

#### **TOPICS OF DISCUSSION** 5

Another way to look at hashtags is looking at the videos they appear in, and which other video they are close to, in order to find clusters of them and pull away big topics of discussions in TikToks.

### 5.1 Methods

The videos were modelled by the list of hashtags they contain. To find topics of discussion, the videos were first clustered by similarity of hashtags present, and then the top frequent-words were analysed to constitute topics.

The hashtags were first vectorised using the TF-IDF method. From that, the pairwise distance between all the Tiktoks was calculated, using cosine similarity. A weighted graph was created, where the nodes were the Tiktoks and the weight of the edges were the distances between two nodes (Tiktoks). The graph was then partitioned using the *community*<sup>2</sup> library, that uses Louvain heuristics to try to find the partition that maximises the modularity [2]. The words that appear the most frequently in each cluster were then assessed, constituting topics that were manually given a name.

<sup>&</sup>lt;sup>2</sup>https://python-louvain.readthedocs.io/en/latest/api.html

### 5.2 Results

The partition of the graph resulted in 21 different clusters, and the graph has a modularity score of 0.646. Modularity measures the relative density of edges inside communities with respect to edges outside communities, and ranges from -0.5 to 1 [3]. The closer to one, the best the partition is.

The list of resulting top-words of each cluster with their manually annotated topic names is shown in Appendix C. The results were aggregated into three different categories: Tiktoks that were general about the election, and contain hashtags from the Biden and Trump sides (Table 3), Tiktoks that were more specific with hashtags about precise issues (Table 4), and finally Tiktoks with hashtags that reflect the young audience nature of Tiktok (Table 5).

### 5.3 Discussion

The graph partitioning has reached a high modularity score, allowing to confidently look at the three main topics present on Tiktok which are discussed below.

5.3.1 Hashtags to reach people. The first striking observation in the result was the rate at which an hashtag appears with its 'opposite' (e.g. *#biden2020* and *#trump2020*). In Table 3, there is not a single topic that does not contain at least one such pair. This is a reflection of how the platform is adopted by the users. Indeed, Tiktok's algorithm chooses which videos the users are going to see, so content creators try to understand it and tweak it to their advantage, in order to reach a wider audience. Hashtags on Tiktok are not tags describing the content, but rather tools to reach the targeted audience. In the context of an election, reaching a wide audience is one of the determining factor in the outcome of the election, and Tiktok users have understood it and use hashtags to have higher chances of appearing on other user's *For You Page*.

5.3.2 Specific topics. Moving on, in the next set of topics emerge ideas of what is discussed on the platform. For example, in Table 4, the topic of Black Lives Matter (BLM) is clearly present, along with other intersectional topics (*#lgbt*). The discussions and events in real life are unsurprisingly reflected on Tiktok. These are events that affect the results of the election, as for example around a fifth of the voters found the BLM protests to be the single most important decision factor in the ballot bow [13]. Tiktok users, by using hashtags and talking about these specific topics, actively participate in the propagation of the debates around the elections.

5.3.3 Young audience. As stated in section 1, the Tiktok audience is mainly young. This was reflected in the last group of topics, with the predominance of hashtags like *#funny, #comedy, #trendy*. These are blueprints of all the memes and jokes made around the election. In a similar way to the point of 5.3.1, the jokes are means found by users to go viral, but in a more powerful way because the political message is passed on implicitly. In an apriori non-political way, the message is disguised. Another interesting result of Table 5 is the presence of an astrology theme, which reveals how the data collection has managed to deviate a bit from the actual election topic.

# 6 CONCLUSIONS

In this work, the political communication on TikTok in the context of the 2020 US presidential election was studied. Concentrating on the set of balanced hashtags representing both Joe Biden and Donald Trump along with few general hashtags, a collection of videos were obtained from TikTok that had at least one of those predefined set of hashtags. The analysis of hashtag appearances by day indicates that momentum to the election was gained steadily till the election day and falls rapidly thereafter. In process of reaching the peak on the election day, the usage pattern showed local peaks on the days of major political events such as the presidential debate and appointment of a supreme court judge. Another important outcome of this analysis was that the hashtags related to Trump occupied the majority of the space at the start of the year, but Biden related hashtags increase their presence only after April 2020 when he became the Democratic Party's presumptive nominee. In trying to understand hashtag usage, the co-occurrences of hashtags were studied. However, the results were surprising and showed that there was no clear demarcation of hashtags between the two political sides. However, a difference in the set hashtags based on their lifetime was seen using community detection. While two communities hashtags were found, the first one has hashtags relevant at that point in time. The second one has hashtags that have meaning and exist for a long time before they fade out. Lastly, graph partitioning for topic modelling was used to identify the topics of discourse. The result of this last analysis showed that the users had used hashtags to reach a wider audience rather than conveying a message through them. Also, they have used hashtags to indicate the social issues in the video in a fun and trendy way. While many users of TikTok do not have the legal age to vote, it appears that they make political videos as well. Notably, Gen Z might not be interested in the politics that is about parties or candidates but they do want to talk about social issues. However, they chose a colloquial way that blends the message and fun. This behaviour was indicated in the way the social hashtags were used.

# 7 LIMITATIONS & FUTURE DIRECTIONS

This article investigated the discourse of the 2020 US presidential elections in TikTok using hashtags. However, this work did not take into account the popularity of the videos. All videos were given the same importance irrespective of their views, likes and shares. These popularity indicators could be included to understand how the videos transmit across users and see if some users act as hubs and others follow these hubs to make their videos. Additionally, although the hashtags serve as a base to convey the idea of TikTok they are associated with, they do not inform what is in TikTok itself. As TikTok is primarily a video-based app, the results of this work can be strengthened by analysing the videos themselves. Two majors paths in this direction would be the speech analysis and the text analysis that appears along with the video.

## REFERENCES

- [1] BBC. 2020. Coronavirus: The world in lockdown in maps and charts. https://www.bbc.com/news/world-52103747
- [2] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* 2008, 10 (oct 2008), P10008. https: //doi.org/10.1088/1742-5468/2008/10/p10008

The US 2020 elections through the eyes of TikTok

- [3] U. Brandes, D. Delling, M. Gaertler, R. Gorke, M. Hoefer, Z. Nikoloski, and D. Wagner. 2008. On Modularity Clustering. *IEEE Transactions on Knowledge and Data Engineering* 20, 2 (2008), 172–188. https://doi.org/10.1109/TKDE.2007. 190689
- [4] Craig Chapple. 2020. TikTok Crosses 2 Billion Downloads After Best Quarter For Any App Ever. https://sensortower.com/blog/tiktok-downloads-2-billion
- [5] Aaron Clauset, M. E. J. Newman, and Cristopher Moore. 2004. Finding community structure in very large networks. *Phys. Rev. E* 70 (Dec 2004), 066111. Issue 6. https://doi.org/10.1103/PhysRevE.70.066111
- [6] Simon Kemp. 2021. Digital 2021 April Statshot Report. https://datareportal.com/ reports/digital-2021-april-global-statshot
- [7] Salzmann Mathieu. 2020. Machine Learning for DH, Lecture 5. (2020).
- [8] Fazal Rizvi, Michael A. Peters, Michalinos Zembylas, and et al. 2020. The long read: On the global relevance of the US elections. *Educational Philosophy and Theory* 0, 0 (2020), 1–20. https://doi.org/10.1080/00131857.2020.1824784 arXiv:https://doi.org/10.1080/00131857.2020.1824784
- [9] Alex Sherman. 2020. TikTok reveals detailed user numbers for the first time. https://www.cnbc.com/2020/08/24/tiktok-reveals-us-global-user-growthnumbers-for-first-time.html
- [10] Aaron Smith. 2009. The Internet's Role in Campaign 2008.
- [11] Sebastian Stier, Arnim Bleier, Haiko Lietz, and Markus Strohmaier. 2018. Election Campaigning on Social Media: Politicians, Audiences, and the Mediation of Political Communication on Facebook and Twitter. *Political Communication* 35, 1 (2018), 50–74. https://doi.org/10.1080/10584609.2017.1334728
- [12] H. Tankovska. 2021. U.S. TikTok users by age 2021. https://www.statista.com/ statistics/1095186/tiktok-us-users-age/
- [13] Sabrina Tavernise and John Eligon. 2020. Voters Say Black Lives Matter Protests Were Important. They Disagree On Why. https://www.nytimes.com/2020/11/ 07/us/black-lives-matter-protests.html
- [14] Joe Tidy and Sophia Smith Galer. 2020. TikTok: The story of a social media giant. https://www.bbc.com/news/technology-53640724

# A LINKS

Github link: https://github.com/JanMaxime/ComputationalSocialMedia\_TikTok/

# **B** LIST OF HASHTAGS FOR DATA COLLECTION

#biden2020, #trump2020, #biden, #trump, #election2020, #2020election, #donaldtrump, #joebiden, #maga, #trumpout, #democrat, #republican, #trumpvsbiden, #bidenvstrump, #voteblue, #votered

# C RESULTS TOPIC MODELLING

# C.1 The two parties in the election

Topic name	Top Hashtags
election day	trumpvs biden, election 2020, greens creen, biden 2020, viral, trump, dumptrump 2020, vote, biden, election and the second sec
election	biden 2020, trump 2020, trump vs biden, trump, election 2020, politics, biden, greenscreen, voteblue, voteblue, trump, election 2020, politics, biden, greenscreen, voteblue,
parties	democrat, republican, liberal, biden 2020, trump 2020, politics, conservative, trump, biden, maga and the set of the se
pro-Trump	trump 2020, vote red, trump, republican, conservative, maga, biden 2020, stitch, kag, politics and the state of the stat
election	biden vstrump, biden 2020, biden, election, trump, trump 2020, 2020 election, election 2020, vote, green screen and the strump
pro-Trump	$trump 2020, maga, votered, trump train, trump, trumps {\tt quad 2020, kag, biden 2020, biden, republican}$
debate	joe biden, donald trump, trump, biden 2020, biden, trump 2020, election, kamala harris, debate, trump vs biden and trump vs b
pro-Biden	vote blue, biden 2020, vote, dumptrump, biden harris 2020, vote him out, trump, trump 2020, biden, politics
election	election 2020, election, biden 2020, trump, biden, trump 2020, vote, 2020 election, trump vs biden, joe biden and trump vs biden and true vs biden and t
Table 3: Topics general about the election	

# C.2 Specific issues discussed

Topic name	Top Hashtags
BLM	blm, black lives matter, biden 2020, trump 2020, democrat, trump, maga, vote blue, lgbt, liberal black lives and the state of the sta
Black pro-Trump	black conservative, make the switch, open your eyes, republican, conservative, trump 2020, democrat, maga, black republican, conservative, trump 2020, democrat, maga, conservativ
democracy	trumpout, biden 2020, usa, trump 2020, biden, trump, tiktok usa, america, use lection, democracy trumpout, biden and trumpout, biden and trump and trumpout, biden and trump and trump and trumpout, biden and trump and true and trump and true and t
Home Office	2020 election, biden 2020, trump, trump 2020, green screen video, vote, election, biden harris 2020, home office, joe biden
democracy (es)	democracia

Table 4: Topics related to specific subjects

# C.3 Topics from young users

CSM'21, Spring, 2021, EPFL

The US 2020 elections through the eyes of  $\mathsf{TikTok}$ 

Topic name	Top Hashtags
Trump	donald trump, trump, usa, funny, viral, america, comedy, trump 2020, melania trump, president trump, trum
Comedy	trump, politics, comedy, election, liberal, republican, progressive, biden, biden 2020, satired to the state of the stat
Astrology	2020 election, zodiac signs, a strology, boo, biden 2020, day in mylife, well done, wip, bbmas, voteblue 2020 to the strong st
Viral	viral, trump 2020, trump, xyzbca, trending, republican, biden, funny, biden 2020, conservative and the second structure of t
Key-words	trump, biden, trumpvs biden, viral, greens creen, biden 2020, voteblue, 2020 election, xyzbca, election 2020 election, where the trump of trump of the true of the t
Sport	trumpout, trump, biden 2020, duet, biden, billy blanks jr, viral, election 2020, vote blue, trumpout 2020 with the second state of the second st
Holiday	holiday tik to k, rnbvibes, election, election 2020, biden 2020, biden vstrump, trump 2020, trump vsbiden, vote, 2020 election and the state of th

Table 5: Topics related to young audience