Analyzing European Migrant-related Twitter Deliberations

Aparup Khatua
L3S Research Center
Leibniz Universität Hannover
Germany
khatua@l3s.de

Wolfgang Nejdl
L3S Research Center
Leibniz Universität Hannover
Germany
nejdl@l3s.de

ABSTRACT

Machine-driven topic identification of online contents is a prevalent task in the natural language processing (NLP) domain. Social media deliberation reflects society's opinion, and a structured analysis of these contents allows us to decipher the same. We employ an NLP-based approach for investigating migration-related Twitter discussions. Besides traditional deep learning-based models, we have also considered pre-trained transformer-based models for analyzing our corpus. We have successfully classified multiple strands of public opinion related to European migrants. Finally, we use 'BertViz' to visually explore the interpretability of better performing transformer-based models.

CCS CONCEPTS

• **Information System**; • **World Wide Web**; • **Information retrieval**;

KEYWORDS

Migration, Twitter, BERT, RoBERTa, BertViz

ACM Reference Format:
Aparup Khatua and Wolfgang Nejdl. 2021. Analyzing European Migrant-related Twitter Deliberations. In *Companion Proceedings of the Web Conference 2021 (WWW '21 Companion), April 19–23, 2021, Ljubljana, Slovenia.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3442442.3453459

1 INTRODUCTION

According to the World Migration Report 2020, more than 50% of all international migrants live in Europe and the USA [17]. Europe's migrant issues are slightly different in comparison to other nations. For instance, migrants in the USA are mostly from their neighboring countries. On the other hand, European migrants (or refugees) are mostly coming from middle east countries[18]. Some refugees take the sea route to enter the European continent through Greece, and then they move from one European nation to another [2, 8]. According to the Eurostat report, 22.3 million people in the European Union (EU) are not citizens of the EU on 1 January 2018, which is 4.4% of the EU population.

Extant literature argued that social media platforms, such as Twitter, can be 'a proxy for reality' in the context of refugees and migrants [2]. Hence, prior studies probed online contents. [10] observe that 'actual situation would influence the intensity and polarity of discussions' on the social meida platforms, and subsequently, it also captures 'how the discussion on refugees changed over time'. Our literature review reveals that existing literature has mostly analyzed social media contents from the perspective of a particular event or issue, but not analyzed the wide range of latent themes. Hence, this paper is attempting to address this gap by identifying multiple strands of public discourse about migrants on Twitter platform in the European context. Classification-based approach of online contents is a prevalent Natural Language Processing (NLP) task [7]. This allows to identify and classify voluminous, but unstructured online contents in a structured manner for policylevel interventions.

Following prior studies, we have employed deep neural models, such as convolutional neural networks (CNN) and Long short-term memory (LSTM), for our initial analysis [12, 29]. We have also employed state-of-the-art pre-trained transformer-based models, such as Bidirectional Encoder Representations from Transformers (BERT) and Robustly Optimized BERT Pretraining Approach (RoBERTa), for our classification task [6, 15]. During the training, we explore different combinations of model hyperparameters to increase our classification accuracies, and this enhances our findings' robustness. We find that BERT and RoBERTa have outperformed CNN and LSTM models. We also find that hyperparameters adjustments can slightly improve the classification accuracies on unseen testing data. Next, we employed BertViz to explore these transformer-based models' interpretability - especially we probed the role of individual layers and attention heads to gauge these model's sensitivity. BertViz helped us to understand why transformer-based models perform better. In brief, our paper has demonstrated the feasibility of NLP-based approaches for analyzing multiple strands of public discourse in the context of the European migration and joins the artificial intelligence (AI) for social good stream of literature.

2 SOCIAL MEDIA AND MIGRATION: A REVIEW

Prior studies have explored Twitter [2, 3, 8, 18, 19], Facebook [4, 11], YouTube [14], and Instagram [9] for probing migration-related issues. One stream of studies probed specific events such as the Paris attack, sexual assault in Cologne, EU-turkey deal for refugees, and other unfortuanate events [21, 23, 24]. For instance, [24] explored Twitter network formation by analyzing various hashtags related to politics, place, and media personalities; whereas [21] considered English and German tweets, and compared the positive and negative emotions, such as anger and anxiety. Similarly, [23] analyzed the

WWW '21 Companion, April 19–23, 2021, Ljubljana, Slovenia

ACM ISBN 978-1-4503-8313-4/21/04.
https://doi.org/10.1145/3442442.3453459

mainstream print media reaction as a response to two unfortunate events, i.e., the death of Alan Kurdi (in 2015) and Omran Daqneesh (in 2016). This study probed whether an unfortunate event can lead to solidarity movement or not. *MMoveT15* Twitter dataset was used to predict migration movement in the context of Germany, Hungary, and Austria [25]. Another study in the context of Organisation for Economic Co-operation and Development (OCED) analyzed geolocated tweets to probe the relationships between internal and international migration [30]. Some of the prior studies also employed content analysis approach. For instance, [20] considered 541 German tweets and analyzed the nature of offensive languages towards refugees by using a 6-point scale. Similarly, [20] explored Turkish hashtags against Syrian refugees and tried to understand the temporal dimensions during July 2016. They found that the hashtag *#IdontwantSyriansinmycountry* was tweeted and retweeted over 54,000 times in a single day. [13] analyzed a small corpus of 100 tweets, and explored how a particular hashtag, i.e., *#refugeesnotwelcome*, was used to express negative feelings, beliefs, and ideologies toward refugees and migrants in Europe. Similarly, [22] explored the social media discourse and mainstream media to understand the perceptions about male refugees coming form middle east. This study argues that certain perceptions emerged due to insignificant presence of female and children refugees from the middle east. A few other studies also analyzed social media data to investigate focused issues, such as deserving vis-à-vis undeserving migrants or security-related concerns vis-à-vis humanitarian-related concerns or critical vis-à-vis positive tweets [9, 10, 19]. [13] argues that nationalist conservative and xenophobic political parties establish a socially accepted racism discourse in the European context. Prior studies also observed that migration issues could influence the electoral discourse in countries like Spain [1, 3] and Italy [4].

To sum up, prior studies have mostly considerd either a specific event [21, 23, 24] or an issue [13, 22, 25, 30]. However, none of the prior studies, to the best of our knowledge, tried to analyze the entire range of migration-related latent issues. Hence, our study is proposing a NLP-based approach for classifying the unstructured Twitter corpus. This will allow policymakers to analyze the diverse range of dominant migrant-related opinions on Twitter platform.

3 TWITTER DATA

We have used the Twitter search API for collecting Twitter data from May 2020 to September 2020. We have considered a set of migrant-related keywords (i.e., UNCHR, asylum, deport, etc.) for our initial crawling. This has yielded around 1.8 million tweets. We have discarded duplicate tweet-ids as well as removed tweets with similar text. Our initial analysis suggests that very short tweets are not insightful. Hence, we have selected English long tweets (i.e., more than 90 characters). Next, we have considered tweets that have specifically mentioned either of these four keywords: 'migrants', 'refugee', 'immigrant', and 'immigration'. This has helped us to discard junk tweets and reduced our corpus to 0.3 million tweets.

We find a significant portion of these 0.3 million tweets are related to the USA context and revolves around Trump's policy-related concerns. Hence, these tweets are not relevant to the European context. To tackle this issue, we have considered migration-related tweets that are specific to the European context. We have checked the linguistic content of tweets and considered tweets, which have categorically mentioned keywords such as 'European Union', 'Europe', 'Germany', 'Greece', 'France', and other European countries. This has resulted in a corpus of around 24,000 unique long English European migration-related tweets. We find that the volume of English tweets is significantly higher than other major European languages such as French, Spanish, and German. Hence, we did not go for a multilingual analysis.

Table 1: Representative Tweets from our corpus

Dominant issues	**Representative tweets** *(678 tweets)*
Safety Concerns *(142 tweets)*	An Afghan migrant and his two sons have been arrested in the brutal stabbing murder of a man aboard a bus in Kiruna, Sweden, according to reports
Economic Conditions *(117 tweets)*	This is shocking - poor conditions of migrant #strawberry pickers in Spain - it is critical those who harvest our food are paid a fair living wage @FFC_Commission
Employment Opportunities *(89 tweets)*	For all #migrant #entrepreneurs & migrant-led organisations that support migrant and #refugee entrepreneurs in Europe to set up their business: WE NEED YOU. The #EMEN-project is building a list of such programmes Complete this FORM by . . .
Healthcare Support *(154 tweets)*	Sara is an Iranian kid who is been living in Turkey as a refugee with her family since 2015. She is suffering from a rare autoimmune disease called Evans Lupus syndrome. Sara can't get medical care she needs in Turkey , please help her and her family @ICHRI @SaveSaraLife
Inequality & Discrimination *(176 tweets)*	The media generate high levels of anxiety about immigration, resulting in negative migrant stereotypes. The public debate in EU has been influenced by populist politicians and biased media coverage with a background of subconscious postcolonial legacy

Next, we refer to the World Migration Report 2020 to understand various concerns, and we tried to map these issues with our tweet corpus. We have classified our tweet corpus into 5 categories: safety concerns; economic conditions; employment opportunities; healthcare support; and inequality & discrimination. For instance, we have considered bigotry, societal inequity, hatred, and prejudice towards migrants as discrimination-related concerns. Similarly, we have considered violence by migrants, and subsequent police action such as arrest, as safety concerns. Table 1 reports a few sample tweets for all these themes. We have tried to take a balanced sample, and considered 678 annotated tweets for our final analysis.

4 METHODOLOGY

Prior studies found that CNN and LSTM models are superior to the traditional bag-of-words approach for sentence classification [5, 12].

Table 2: Accuracies of CNN and LSTM models

CNN	LSTM	Batch Size	Dropout Rate
0.8319	0.7058	16	0.3
0.7341	0.6985	16	0.5
0.8088	0.5955	32	0.3
0.6029	0.7058	32	0.5

Hence, we have employed CNN and LSTM as our baseline models. We have also considered two transformer-based models: BERT [6]and RoBERTa [15]. BERT is a bidirectional unsupervised pre-trained model that considered BooksCorpus and English Wikipedia (16GB) for the training purpose [6]. RoBERTa has considered an additional corpus of CC-News (76 GB), Open Web Text (38 GB), and Storie's dataset (31 GB) for training [15]. This huge training corpus allowed these bidirectional transformers models to effectively interpret the contextual meaning. We have considered the HuggingFace python library [28] which not only includes pre-trained models but also allows fine-tuning of hyperparameters. For our classification task, we have employed *'BertForSequenceClassification'*. Subsequently, we have considered the BERT-Base-Uncased model comprised of 12-layers and 12-heads with a total of 110M parameters.

5 FINDINGS

Table 2 reports the classification accuracies of CNN and LSTM models for multiple batch sizes (16 and 32) and dropout rates (0.3 and 0.5). The hidden layer for all these models was 256. We have considered SoftMax activation in our final classification layer to predict the final class. We use 'rmsprop' as our optimizer. Accuracies were broadly in a similar range for various combinations of batch sizes and dropout rates. Classification accuracies for CNN models are mostly higher than LSTM models, and the highest accuracy is 83.2%.

Table 3 reports the findings of BERT and RoBERTa models. [6] recommend to jointly optimize the sequence length and max batch size to tackle out-of-memory issues. We employed Adam optimizer, which requires a significant computational resource. Hence, we have done our analysis using Google Collaborative Environment (Colab). Details of our computation resource were as follows: 25.51 GB RAM along with a Tesla P100-PCIE-16GB GPU. We have considered the following combinations: batch sizes (16, 32) and learning rates (2e-5, 3e-5, 5e-5). These combinations result in a total of 6 varieties for each of our transformer-based models. We have considered a maximum sequence length of 256, which is sufficient for our tweet corpus. We have used 4 epochs for our BERT-Base-Uncased models. Our open-source implementation, pre-trained weights, and full hyperparameter values are in accordance with the HuggingFace transformer library [28].

The accuracies of transformer-based models are better than our base model, and we find accuracies are mostly more than 90%. This is a significant improvement with respect to base models. We also observe that RoBERTa models have perfomed marginally better than BERT models. Next section visually explores how contextual interpretation of transformer-based models has improved the performance.

Table 3: Accuracies of BERT and RoBERTa models

BERT	RoBERTa	Batch Size	Learning Rate
0.9144	**0.9265**	16	2e-5
0.8877	0.9191	16	3e-5
0.9037	0.9191	16	5e-5
0.9037	**0.9265**	32	2e-5
0.9091	0.8676	32	3e-5
0.9091	0.8750	32	5e-5

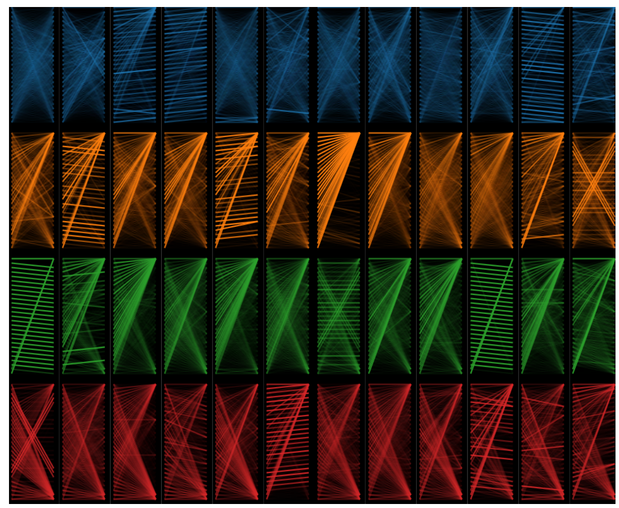

Figure 1: Model view of BERT for a sample tweet *(excludes layers 4 to 11)*

6 INTERPRETABILITY OF BERT AND ROBERTA

Following [26, 27], we visually explored pre-trained models' attention weights using the BertViz library. This library is an improvement of the original tensor2tensor based implementation of [16]. BertViz allows us to interpret and analyze the multi-head self-attention weights of the tokens from different BERT and RoBERTa model layers. BertViz tool can also generate the overall model view, which helps us to see the 'attention across all of the model's layers and heads for a particular input' tweet [26, 27]. All attention heads are reported in a matrix format in this model view, where rows represent the layers and columns represent the heads. [26] pointed out this snapshot view captures the contextual relation between tokens for all 12 layers and all heads. Figure 1 reports the model view for a tweet as follows: *'If you haven't realized that nationalism and antiimmigrant violence is alarmingly in Europe then you haven't been paying attention'.*

Reporting the entire 12*12 matrix will make the graphics clumsy. Hence, we have reported a truncated version of 12*4 (i.e., 0 to 11

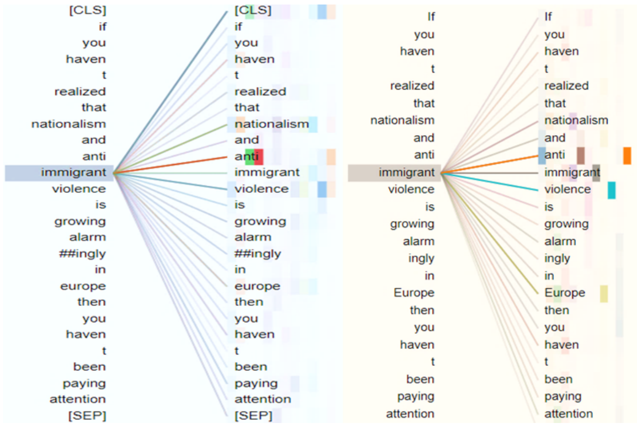

Figure 2: BERT (left) and RoBERTa (right) attention to the word *'immigrant.'*

heads and 0-3 layers). Careful observation reveals a few horizontal-stripe patterns (e.g., layer 0 & heads 2/3/10; layer 2 & heads 0/6/9) and a few triangular patterns (e.g., most heads in layer 1). This is similar to prior studies. [26] argues that a horizontal-stripe indicates that 'tokens attend to the current position, whereas a triangular pattern indicates that 'they attend to the first token'. However, this overall model view does not allow to probe the exact contextual relationship between tokens.

Hence, we probe the attention-head view that allows us to interpret the attention of one or more heads. In this graphical interface, self-attention is represented as lines from left to right, and the thickness of these lines represent the attention weight between tokens. However, some of the attention weights between tokens can be very low, and they will be nearly invisible. Different colors used in this visualization represent the individual heads in the model. This attention-head view enables us to explore the precise attention heads of our input sentences. An in-depth analysis of these attention weights between tokens reveals how models, such as BERT and RoBERTa in our case, interpret the underlying relationship between words/tokens by allocating more attention to particular parts of the input text. Subsequently, these higher attentions, i.e., thicker lines, play a crucial role in the downstream task, which is classification in our case.

Figure 2 reports the functioning of BERT (the left figure) and RoBERTa (the right figure) models for the previous tweet of Figure 1. This figure indicates why RoBERTa models have performed slightly better than BERT models for the classification task. For instance, the BERT model links the token *'immigrant'* with the token *'anti'*. However, in addition to this *'anti'* token, the RoBERTa is also linking *'immigrant'* with *'violence'* and *'Europe'*. Thus, the RoBERTa model is more appropriately capturing the *European* context. Additionally, the model is also linking the issue 'immigrant' with the token 'anti' and 'violence' that captures the overall apprehensive sentiment of the original tweet. In short, BertViz helps us to visualize how transformer-based models assign different weights to different tokens based on their training and outperform other models.

7 CONCLUSION

A plethora of studies have explored social media data, especially Twitter platform, for investigating public opinions across a wide range of socio-economic issues. A handful of studies also analyzed Twitter data in the context of migration. However, most of these studies probed a specific event or analyzed a specific issue. To the best of our knowledge, none of the prior studies tried to identify and classify a diverse range of concerns on social media platform. Our study has made a humble effort in this direction. We do acknowledge that Twitter has some inherent limitations. First, it is voluminous but unstructured. Second, this text data allows to do opinion mining, but only a minuscule portion of the society is vocal on the Twitter platform. Addressing this second concern is beyond the scope of our study. However, our study has addressed the first limitation by classifying the unstructured corpus into issue-specific dominant migration topics (which can be captured through social media). We also showed, by using BertViz graphics, why transformer-based models are better suited to capture these diverse issues. However, future studies need to go beyond this and perform a more fine-grained analysis for policy-level interventions. For instance, the information about an employment opportunity should not get lost, but it should reach a job-seeking migrant. NLP researchers need to use data science for social good. Our study took a baby step in this direction in the context of European migration.

ACKNOWLEDGMENTS

Funding for this project was, in part, provided by the European Union's Horizon 2020 research and innovation program under Grant Agreement No: 832921.

REFERENCES

[1] Alcántara-Plá, M. and Ruiz-Sánchez, A. 2018. Not for Twitter: Migration as a Silenced Topic in the 2015 Spanish General Election. *Exploring Silence and Absence in Discourse.* M. Schröter and C. Taylor, eds. Springer International Publishing. 25–64.
[2] Aswad, F. and Menezes, R. Refugee and Immigration: Twitter as a Proxy for Reality. 6.
[3] Calderón, C.A. *et al.* 2020. Topic Modeling and Characterization of Hate Speech against Immigrants on Twitter around the Emergence of a Far-Right Party in Spain. *Social Sciences.* 9, 11 (Oct. 2020), 188. DOI:https://doi.org/10.3390/socsci9110188.
[4] Capozzi, A. *et al.* 2020. Facebook Ads: Politics of Migration in Italy. In: Aref S. *et al.* (eds) Social Informatics. SocInfo 2020. Lecture Notes in Computer Science, vol 12467. Springer, Cham. https://doi.org/10.1007/978-3-030-60975-7_4.
[5] Conneau, A. *et al.* 2017. Very Deep Convolutional Networks for Text Classification. Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers (Valencia, Spain, 2017), 1107–1116.
[6] Devlin, J. *et al.* 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv:1810.04805 [cs].* (May 2019).
[7] Dilrukshi, I. and Zoysa, K. de 2014. A Feature Selection Method for Twitter News Classification. *International Journal of Machine Learning and Computing.* 4, 4 (2014), 365–370. DOI:https://doi.org/10.7763/IJMLC.2014.V4.438.
[8] Gualda, E. and Rebollo, C. 2016. The Refugee Crisis On Twitter: A Diversity Of Discourses At A European Crossroads. 3 (2016), 14.
[9] Guidry, J.P.D. *et al.* 2018. Welcome or Not: Comparing #Refugee Posts on Instagram and Pinterest. *American Behavioral Scientist.* 62, 4 (Apr. 2018), 512–531. DOI:https://doi.org/10.1177/0002764218760369.
[10] Hadgu, A.T. *et al.* 2016. Welcome or Not-Welcome: Reactions to Refugee Situation on Social Media. *arXiv:1610.02358 [cs].* (Oct. 2016).
[11] Hrdina, M. 2016. Identity, Activism and Hatred: Hate Speech against Migrants on Facebook in the Czech Republic in 2015. *Naše společnost.* 1, 14 (2016), 38. DOI:https://doi.org/10.13060/1214438X.2016.1.14.265.
[12] Kalchbrenner, N. *et al.* 2014. A Convolutional Neural Network for Modelling Sentences. *arXiv:1404.2188 [cs].* (Apr. 2014).

[13] Kreis, R. 2017. #refugeesnotwelcome: Anti-refugee discourse on Twitter. *Discourse & Communication*. 11, 5 (Oct. 2017), 498–514. DOI:https://doi.org/10.1177/1750481317714121.

[14] Lee, J.-S. and Nerghes, A. 2018. Refugee or Migrant Crisis? Labels, Perceived Agency, and Sentiment Polarity in Online Discussions. *Social Media + Society*. 4, 3 (Jul. 2018), 205630511878563. DOI:https://doi.org/10.1177/2056305118785638.

[15] Liu, Y. *et al.* 2019. RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv:1907.11692 [cs]*. (Jul. 2019).

[16] Llion, J. 2017. Tensor2tensor transformer visualization. https://github.com/ tensorflow/tensor2tensor/tree/ master/tensor2tensor/visualization.

[17] McAuliffe, M., James *et al.* 2019. *World migration report 2020*.

[18] Nerghes, A. and Lee, J.-S. 2019. Narratives of the Refugee Crisis: A Comparative Study of Mainstream-Media and Twitter. *Media and Communication*. 7, 2 (Jun. 2019), 275–288. DOI:https://doi.org/10.17645/mac.v7i2.1983.

[19] Nerghes, A. and Lee, J.-S. 2018. The Refugee/Migrant Crisis Dichotomy on Twitter: A Network and Sentiment Perspective. *Proceedings of the 10th ACM Conference on Web Science* (Amsterdam Netherlands, May 2018), 271–280.

[20] Özerim, M.G. and Tolay, J. 2020. Discussing the populist features of anti-refugee discourses on social media: an anti-Syrian hashtag in Turkish Twitter. *Journal of Refugee Studies*. (May 2020), feaa022. DOI:https://doi.org/10.1093/jrs/feaa022.

[21] Pope, D. and Griffith, J. 2016. An Analysis of Online Twitter Sentiment Surrounding the European Refugee Crisis: *Proceedings of the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management* (Porto, Portugal, 2016), 299–306.

[22] Rettberg, J.W. and Gajjala, R. 2016. Terrorists or cowards: negative portrayals of male Syrian refugees in social media. *Feminist Media Studies*. 16, 1 (Jan. 2016), 178–181. DOI:https://doi.org/10.1080/14680777.2016.1120493.

[23] Sajir, Z. 2019. Solidarity, Social Media, and the "Refugee Crisis": Engagement Beyond Affect. (2019), 28.

[24] Siapera, E. *et al.* 2018. Refugees and Network Publics on Twitter: Networked Framing, Affect, and Capture. *Social Media + Society*. 4, 1 (Jan. 2018), 205630511876443. DOI:https://doi.org/10.1177/2056305118764437.

[25] Urchs, S. *et al.* 2019. MMoveT15: A Twitter Dataset for Extracting and Analysing Migration-Movement Data of the European Migration Crisis 2015. *2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)* (Napoli, Italy, Jun. 2019), 146–149.

[26] Vig, J. 2019. A Multiscale Visualization of Attention in the Transformer Model. *arXiv:1906.05714 [cs]*. (Jun. 2019).

[27] Vig, J. and Belinkov, Y. 2019. Analyzing the Structure of Attention in a Transformer Language Model. *arXiv:1906.04284 [cs, stat]*. (Jun. 2019).

[28] Wolf, T. *et al.* 2020. HuggingFace's Transformers: State-of-the-art Natural Language Processing. *arXiv:1910.03771 [cs]*. (Jul. 2020).

[29] Young, T. *et al.* Recent Trends in Deep Learning Based Natural Language Processing. 23.

[30] Zagheni, E. *et al.* 2014. Inferring international and internal migration patterns from Twitter data. *Proceedings of the 23rd International Conference on World Wide Web - WWW '14 Companion* (Seoul, Korea, 2014), 439–444.