Topic Modelling and Emotion Detection on Italian Tweets during the early Covid-19 Pandemic

Bachelor's Project Thesis

Malina Chichirau, s3412768, m.chichirau@student.rug.nl,
Supervisors: Dr J. Spenader (internal) and Dr. Tomasso Caselli (external)

Abstract: This exploratory study attempts to identify the major topics of discussion in Italian tweets dated between February and July 2020, and the emotions associated with these topics. Topic Modelling was achieved using the SeaNMF algorithm, which exploits word contexts as a proxy for semantic similarities, while emotion detection was performed through lexicon look-up. Results revealed that the distribution of topics was imbalanced and found little evidence for a direct connection between changes in topic trends and the events selected as reference points. By examining the keywords of a given topic over the period of the study, a shift is observed in the focus of some topics. An emotion analysis of the tweets found similar patterns in terms of the intensities and fluctuations of emotions, regardless of the topics the tweets concerned. A comparison between Italian and Dutch tweets collected in the same period indicates that Italians were more preoccupied with internal affairs than their Dutch counterparts.

Keywords: Covid-19; Italy; Twitter; Topic Modelling; SeaNMF; Emotions.

1 Introduction

Italy was the first country in Europe to impose a nation-wide quarantine on March 10th (Legorano and Sylvers, 2020), following reports of over 10 000 cases of infection with the SARS-CoV-2 virus, and over 600 deaths associated with the coronavirus disease 2019 (Dong, Du, and Gardner, 2020). During the quarantine, Italian people spent more time online, which resulted in a 52% increase of the average time spend on social media*, on sites such as Facebook, Twitter, and Instagram. Italian people also spend 67% more time following news coverage*, 74% of them reporting that their primary source of news was an online medium, including social media (Newman, Fletcher, Schulz, Andı, and Nielsen, 2020).

In fact, studies indicate that Twitter was an effective platform for government agencies, media outlets, and international organisations to raise awareness and circulate important information, during the Zika virus (Daughton and Paul, 2019), and the Ebola outbreaks (Househ, 2016). However, the nature of an open social media platform such as Twitter, encourages people not only to engage with the verified news distributed by official sources, but also to express their own opinions and experiences, start new conversations, address different aspects of current events, and in some cases question and undermine authorities, or spread misinformation. Furthermore, the intensity and types of emotions conveyed by information have been shown to influence people's attitudes towards it and their likelihood to believe it (Martel, Pennycook, and Rand, 2019). Therefore, identifying the popular topics of discussion on Twitter and people's emotions towards them, can be both a good indicator of the public opinion regarding matters of national importance, and a first step in preventing the spread of misinformation.

Previous studies which monitored the Twitter activity during various stages of the Covid-19 pandemic have been generally targeting tweets written in English, and mainly concerning the situation in the United States of America (Ordun, Purushotham, and Raff 2020; Xue, Chen, Hu, Chen, Zheng, Su, and Zhu 2020), with few exceptions such as Marinov, Spenader, and Caselli (2020). Marinov et al. (2020) studied the topics discussed and emotions expressed by different sub-groups of Dutch

*https://www.statista.com/statistics/1106498/home-media-consumption-coronavirus-worldwide-by-country/

and Belgian Twitter users. They pointed out that focusing on a smaller community has advantages in terms of observing a more direct link between the development of the pandemic, the actions of authorities and the reactions of the general public. Consequently, the current study attempts a similar endeavour by analyzing tweets written in Italian, between February and July 2020, from a corpus compiled by Basile and Caselli (2020). Furthermore, this study seeks the answers to three research questions.

First, the study inquires whether the topics related to the Covid-19 pandemic and addressed by Italian tweets changed during the months between February and July, and whether the changes were related to notable events which occurred in that time-frame. Second, the study attempts to compare the emotions expressed by Twitter users in relation to each topic, and whether the emotions fluctuated or shifted as a result of notable events. Lastly, as the emotions will be evaluated using a lexicon containing emotion scores for words (Mohammad and Turney, 2013), and a lexicon containing emotion scores for emojis (Shoeb, Raji, and de Melo, 2019), a third research question compares the emotions expressed through words with those expressed through emojis.

2 Methods

This section first introduces a series of events related to the Covid-19 pandemic in Italy (Section 2.1), which will be used as reference points when examining the trends in the topics and emotions addressed by the tweets. Then it presents the 40wita corpus, data collection and pre-preprocessing in Section 2.2. Then, the Topic Modelling technique is introduced, and its limitation for short texts are discussed in Section 2.3. Next, the SeaNMF method is introduced and explained in detail, along with its hyper-parameters and the evaluation metrics used for interpreting the results, in Section 2.4. Finally, Section 2.5 presents the lexicons used for emotion detection, the emotions included in the study and the derived metrics.

2.1 Notable Events

In order to study how Twitter discussions were shaped by the development of the pandemic in Italy, a list of 12 representative events was compiled. The events were selected to reflect the important milestones in the spread of the virus (for instance the confirmation of the 3rd case, and the ultimate decrease in deaths and active cases), the response of the authorities (the start of the lockdown, the closure of non-essential factories), the gradual relaxation of lockdown restrictions (the moment when Italy opened its borders to European tourists), and also the impact of the lockdown on culture and sports (the cancellation of the Venice Carnival, the re-opening of theaters and sports venues). These notable events are presented in Table 2.1, and they are also assigned a label which will be used to reference them in plots.

2.2 Data Pre-processing

The data set used is 40wita (Basile and Caselli, 2020), which resulted by filtering the tweets dated between February 1st, 2020 and July 31st, 2020 from the Twita corpus (Basile, Lai, and Sanguinetti, 2018) on the basis of keywords related to the Covid-19 Pandemic. Examples of keywords used to select tweets include variations of "*covid-19*", "*coronavirus*", "*quarantena*" ("quarantine"), "*stateacasa*" ("stay home") and hashtags associated with social movements "*iononsonounvirus*" ("I am not a virus"), publicity campaigns "*milanononsiferma*" ("Milano does not stop") or mentioning national institutions "*INPSdown*" (National Institute of Social Security). The number of tweets selected each month varies between 163 899 in June and 1 129 703 in March, with a total of 2 896 691 tweets.

In order to identify the main topics in the 40wita corpus, the duplicate tweets were removed. Next, the text of the remaining tweets was pre-processed such that the user-references, tags, hyperlinks, and emojis were removed. The keywords previously used to select the tweets were removed because they are not informative due to their frequency in the corpus; and numeric characters were also removed since when taken out of context, they are not informative either. Stop words were removed using the NLTK Python library (Bird, Klein, and Loper,

Table 2.1: Notable events related to the Covid-19 pandemic which occurred during the period in which the tweets were collected. The labels of the are also used in graphs.

Label	Date	Description of the event
A	February 20	Third confirmed Covid-19 case in the province of Lombardy
B	February 23	The Venice Carnival is cut short by Covid-19 outbreaks
C	March 4	Schools and Universities are closed
D	March 9	A nationwide lockdown is imposed
E	March 22	Nonessential factories are closed
F	March 31	The officials announce that the peak of the pandemic was reached
G	April 5	The number of daily deaths recorder due to Covid-19 starts decreasing
H	April 20	The number of active infections with Covid-19 starts decreasing
I	May 4	Restrictions are relaxed, as Italy enters Phase Two of the lockdown
J	June 15	Theatres, cinemas, sport venues, and playgrounds are opened, as Italy enters Phase Three of the lockdown
K	July 2	Italy opens its borders to European tourists
L	July 14	Nightclubs, fairs and conventions are allowed to open

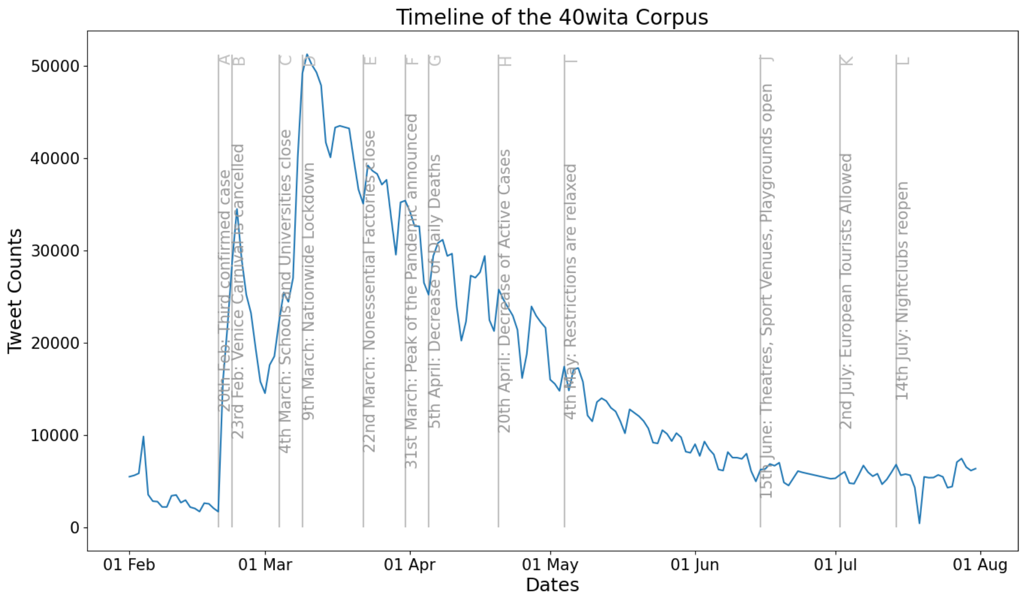

Figure 2.1: Number of tweets collected each day between February 1st and July 31st. Important events which occurred in this time frame are marked by a vertical line and a short description.

2009) and closed-class words are removed by first tagging parts of speech using the SpaCy API (Honnibal and Montani, 2017) and then words tagged as determiner, interjection, numeral, pronoun, adposition, punctuation, symbol, space, auxiliary or any type of conjunction were excluded.

After pre-processing, 2 893 399 tweets remained, with the most tweets being collected during March (1 128 466), and the fewest during June (163 683). The distribution of the collected tweets between February 1st and July 31st can be seen in Figure 2.1, which also indicates notable dates related to the progression of the Covid-19 pandemic in Italy. Intuitively, following an important event, such as the closure of non-essential factories on April 20th, or the cancellation of the Venice Carnival announced on February 23rd, there is a surge in tweets concerning the pandemic. The most distinct spikes are registered in the beginning on the pandemic, and they are related to the aforementioned cancellation of the Venice Carnival and the announcement of the nation-wide lock-down on March 9th, after which the number of tweets about the pandemic peaked on March 10th (over 50 000 tweets).

2.3 Topic Modelling

Topic modelling is an unsupervised machine learning technique used to identify patterns in a corpus. The patterns are clusters of words which co-occur often and describe the topics in the corpus (Brett, 2012), under the assumption that the words in the emerging clusters are semantically related. However, this assumption does not always hold for short texts, because they contain fewer words, thus consistent and meaningful patterns are less likely to occur overall.

Previous research showed that in order to overcome the lack of contextual information short texts can be aggregated to form longer texts, using auxiliary information such as the name of the author, time, and location. For instance, Choo, Lee, Reddy, and Park (2015) aggregated tweets posted by the same author in a pseudo-document. Another method, introduced by Xun, Gopalakrishnan, Ma, Li, Gao, and Zhang (2016), utilizes word a vector representation of words - where semantically related words are represented by vectors projected in the same vector space - in order to infer semantic similarities between words.

However, the aforementioned methods of topic modelling for short texts are both limited by the need for additional resources either in the form of auxiliary information or vector representations trained using external documents. Therefore, the topic modelling method utilized by this paper - SeaNMF (Shi, Kang, Choo, and Reddy, 2018) aims to bypass the need for external resources, by approximating semantic relations using the word contexts provided in the short texts.

2.4 SeaNMF

The original non-negative matrix factorization algorithm - NMF (Lee and Seung, 1999) decomposes a high-dimensional matrix representation of a series of documents in terms of keywords into two lower-dimension matrices which represent the documents in terms of their topics, and the topics in terms of their keywords. Therefore, NMF uses a bag-of-words method to represent the documents, the keywords being selected according to their frequency count in the corpus. Next, a matrix representation of the documents is created using the frequency counts of the keywords in each document. Thus, for N documents and M keywords, their representation is a matrix A of the form $N * M$. The algorithm then iteratively trains two matrices, one being of the form $N*K$ representing the documents in terms of the topics (H), and the other one being of the form $K * M$ representing the topics in terms of keywords (W), where K is the number of topics and it is a hyper-parameter. Consequently, the NMF algorithm offers as output both a list of identified topics and their corresponding clusters of words, and a classification of the input documents by topics.

The SeaNMF algorithm (Shi et al., 2018) improves the NMF algorithm by extending the representation of the documents in terms of keywords with a representation of the contexts in terms of keywords. Contexts consist of the words in a tweet, excluding keywords. Since tweets are already short texts, the span of the contexts is equal to the span of the tweets themselves. To this end, two tweets that differ only by one word - which is a keyword in both tweets - would provide a single context that is linked to both keywords, and each keyword is linked to a different original document.

The A matrix - used by the traditional NMF algorithm - is replaced by a complex matrix S, shown in the left-hand side of Figure 2.2. The values of the matrix S represent the semantic relationships between keywords and their contexts. The computation of the semantic relation is based on the Skip-Gram with Negative-Sampling (SGNS) neural word embedding method proposed by Mikolov, Chen, Corrado, and Dean (2013a; 2013b). Thus, the formula used is the following.

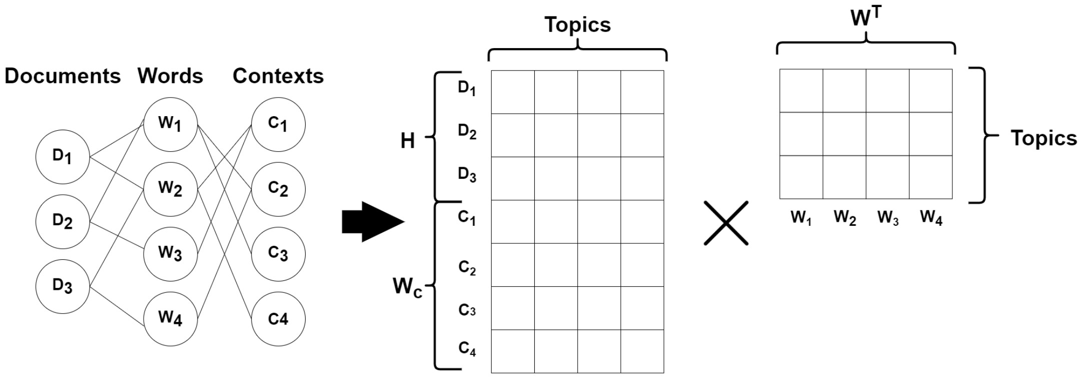

Figure 2.2: Representation of the SeaNMF algorithm (adapted from Shi et al. (2018)). The S matrix the left-hand side of the arrow is the representation of the documents and contexts in terms of keywords; and it is decomposed in a vertical and a horizontal matrix. The vertical matrix is composed of H - the representation of documents in terms of topics, and W_c - the representation of contexts in terms of topics. The horizontal matrix W^T is the representation of topics in terms of words.

$$S_{ij} = \left[log \left(\frac{\# \left(w_i, c_j\right)}{\# \left(w_i\right) * p \left(c_j\right)} \right) - log \left(K\right) \right]_+$$

where S_{ij} is the element of the matrix S in the row i and column j, $\# \left(w_i, c_j\right)$ is the number of occurrences of the word w_i in the context c_j, $\#(w_i)$ is the number of occurrences of the word w_i in the entire corpus, K is the number of topics, and $p \left(c_j\right)$ is the unigram distribution of a context c_j given by the formula:

$$p \left(c_j\right) = \frac{\#(c_j)^\gamma}{\sum_{c_j \in V} \#(c_j)^\gamma}$$

where γ is a smoothing factor, $\#(c_j)$ is the number of occurrences of context c_j. Note that the formula for S_{ij} constrains the value of S_{ij} to be a positive number, which is a requirement of the base NMF algorithm; any negative values are replaced by 0.

Finally, the function which describes the training of the matrices is:

$$\min_{W, W_c, H \geq 0} \left\| \begin{pmatrix} A^T \\ \sqrt{a} S^T \end{pmatrix} - \begin{pmatrix} H \\ \sqrt{a} W_c \end{pmatrix} W^T \right\|_F^2 + \Psi(W, W_c, H)$$

where W_c is the representation of contexts in terms of topics, similarly to how H is the representation of documents in terms of topics (both are illustrated in Figure 2.2) , α is a scaling factor, a hyper-parameter, and Ψ is a penalty function.

2.4.1 Parmeters

The SeaNMF algorithm only requires tuning 3 hyper-parameters α, the scaling factor of the semantic correlation matrix, the number of topics K, and the number of keywords. After trying several values, α was set to 1. This value was in line both with Shi et al. (2018) suggestion for using SeaNMF to model the topics in tweets, and Marinov et al. (2020) choice for tuning α. Furthermore, the number of keywords was set to 10 000 in an attempt to include as many words as possible, but also to balance the time needed for training. Finally, the number of topics was chosen according to the metrics explained in the next section. The choice itself is motivated as part of the Results section.

2.4.2 Evaluation Metrics

In order to evaluate the patterns identified by the topic modelling algorithm, three metrics are used. The PMI and NPMI are measuring the strength of the association between the keywords which are comprising a given topic. Whereas, TD measures how diverse the identified topics are in terms of unique keywords.

The PMI is the pointwise mutual information, which is a measure of association. It compares the probability of two events (x and y) occurring together with the probabilities of them occurring independently, and it is computed using the following formula. The higher the value of PMI, the more closely associated the two events are.

$$PMI = log\left(\frac{P(x,y)}{P(x) * P(y)}\right)$$

The NPMI is the normalized pointwise mutual information metric, which is derived from PMI, but its values are constrained to the interval $[-1, 1]$, with -1 indicating no co-occurrences between the events and 1 indicating that the events are always simultaneous. The formula of NPMI is the following.

$$NPMI = \frac{PMI}{-log\left(P(x,y)\right)} = \frac{log\left(P(x) * P(y)\right)}{log\left(P(x,y)\right)} - 1$$

To values of PMI and NPMI reported in Table 3.1 are computed between the top 10 keywords of every topic. Thus checking how often the keywords identified for each topic co-occur in the documents.

Finally, the TD is the topic diversity metric which compares the number of unique keywords for each topic, relative to the number of topics. It is computed for the top 10 keywords, thus the formula for TD is the following.

$$TD = \frac{\sum_{i=1}^{k} u_i}{10 * K}$$

where u_i is the number of keywords among the 10 top keywords associated with topic i which are not among the top 10 keywords associated with another topic (i.e. they are unique to topic i), and K is the number of topics. The larger the value of TD is, the less overlap there is between the keywords associated with different topics. Thus, a large TD value indicates that there is a greater variation in the top 10 keywords between different topics, suggesting that there is less overlap between the topics. Consequently a large TD value would imply that the identified topics are very distinct, and that a broader range of subjects of discussion have been detected. However, focusing on identifying very distinct topics means that subtopics or topics that are likely to share several keywords (for instance, discussions about Covid-19 infections in Italy and in Europe) will be combined in a single topic. As the purpose of the study is to examine the range of topics related to the pandemic, more distinct topics - thus, a larger TD value - is a preferred result.

2.5 Emotion Detection

In order to fully capture the opinions conveyed by the tweets emotion detection was performed automatically using two lexicons. The Emotion Lexicon EmoLex composed by the National Research Council of Canada (NRC) (Mohammad and Turney, 2013) and the lexicon EmoTag (Shoeb et al., 2019) are both providing ratings for the following emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. The NRC (Mohammad and Turney, 2013) is an automatic translation of the manually composed NRC Emotion Lexicon EmoLex for the English language and it contains 14 182 words. The EmoTag (Shoeb et al., 2019) features the ratings of 150 frequently used emojis on Twitter, with regard to the emotions aforementioned. The scores included in the lexicons are the result of manual annotation of words - in the case of the NRC Emotion Lexicon EmoLex - and emojis - in the case of the EmoTag lexicon.

These specific emotions are used because according to Plutchik (1980) they are the primary emotions which trigger behaviors that favor survival. Furthermore, these emotions represent pairs of opposites (anger-fear, anticipation-surprise, joy-sadness, disgust-trust).

In order to compare the emotion scores yielded by the two lexicons, an emotion score was computed for each tweet, using each lexicon. One of the emotion scores only considered the words in a tweet and it is the sum of the scores of the words in that tweet which are included in the NRC Emotion Lexicon EmoLex. The other score only considered the emojis included in tweets, and was computed by taking the sum of the scores assigned by the EmoTag lexicon to the emojis present in each tweet.

For instance, the following tweet can be translated in English as "@Ministry of Health. So, if a person is infected and does not know it, they will only put on a mask, as a precautionary method, once the first symptoms appear. In the meantime, they have already transmitted the infection". All emojis in this tweet are found to convey various degrees of emotions, and the word *contagio* conveys anticipation, disgust and fear.

> @MinisteroSalute Quindi se una persona è contagiata e non lo sa quando appariranno i primi sintomi come misura precauzionale si metterà la mascherina 😷 Nel frattempo ha già trasmesso il contagio 👈💀

In addition to the emotion scores, a positive and a negative score were computed, as the sums of the 'positive' emotions (joy, trust) and of 'negative' emotions (anger, fear, disgust, sadness). The values of the positive and the negative scores are both positive, since the values of the emotion scores are also positive. Lastly, the polarity score was considered as the difference between the positive and the negative scores. The value of the polarity score can be positive, indicating that the emotions conveyed by a certain tweet are mostly positive, or negative, indicating the opposite. The polarity can be more reliable than the individual emotion scores in the case of automatic emotion detection, since word senses are not always properly disambiguated and thus erroneous emotion scores may be accidentally counted.

3 Results

This section presents the evaluation of the topic modelling technique used to identify the main topics of the 40wita corpus. Next, the frequency of identified topics is plotted against a timeline of notable events which occurred in Italy in the same time-frame, in order to analyse possible correlation between the public discourse and these events. Then, a comparison is made between the emotions associated with each of the main topics identified. Finally, the polarities of the main topics are computed according to each lexicon. Thus, the directionality of the emotions expressed through words and through emojis are compared.

3.1 Topic Modelling Evaluation

Table 3.1 illustrates the values of the metrics presented in Section 2.4.2, computed for different numbers of topics, for each month. All metrics values are computed for the top 10 keywords of every topic. The value in the topics column shows the number of words clusters the SeaNMF algorithm is identifying in the given documents. These values for topics numbers had been also tested by Marinov et al. (2020), and some additional values (< 30 topics) had been tried, in order to improve topic diversity (TD).

For all months, the number of topics seems to be inversely proportional with the topic diversity, with the exception of March and April, where using 30 and 20 topics, respectively lead to equal or greater values for TD than using 10 topics. Thus, the more topics are used, the smaller the TD values - indicating that more of the top keywords which define the topics tend to be shared between the topics. However, the PMI value increases with the number of topics, indicating that there are more co-occurrences of the keywords defining each topic when there are more topics. Therefore, the more topics there are the more coherent the topics seem to be according to PMI values. On the other hand, the NPMI values - computed by normalizing PMI values to account for the difference in frequency of different topics and keywords - are inversely proportional with the number of topics. Furthermore, the NPMI values are slightly negative, indicating that the keywords belonging to the same topic are independent of each other. In fact, the more topics are used, the less likely the keywords within a topic are to co-occur in a tweet, according to NPMI values.

Overall, the metrics suggest that there might be a large difference between the frequency of popular topics, which in turn means that the choice of keywords is not representative of the true range of topics. The explanation behind this observation is that if most of the keywords used for topic modelling belong to a single, very popular and coherent topic, then when trying to identify a larger number of topics those topics will be mostly "subtopics" of the popular topic. These subtopics yield a low TD score - since the top keywords are shared between topics - and also low coherence scores (PMI and NMPI) - since the top keywords are not grouped

Table 3.1: Topic models for each month evaluated according to the metrics PMI, NPMI, TD

	February			March			April			May		
Topics	PMI	NPMI	TD	PMI	NPMI	TD	PMI	NPMI	TD	PMI	NPMI	TD
10	1.552	-0.075	0.54	1.582	-0.073	0.51	2.033	-0.095	0.60	2.085	-0.103	0.68
20	1.869	-0.092	0.48	1.621	-0.074	0.43	2.32	-0.11	0.70	1.966	-0.096	0.57
30	2.531	-0.125	0.45	1.831	-0.082	0.51	2.198	-0.103	0.50	2.474	-0.121	0.52
50	2.721	-0.137	0.32	1.744	-0.078	0.41	2.315	-0.107	0.50	2.513	-0.123	0.51
70	2.988	-0.152	0.36	1.868	-0.083	0.22	2.314	-0.107	0.47	2.763	-0.136	0.43
90	3.047	-0.155	0.33	2.081	-0.094	0.24	2.351	-0.109	0.41	2.706	-0.134	0.34
110	3.149	-0.161	0.32	2.012	-0.091	0.36	2.389	-0.111	0.42	2.816	-0.14	0.36

	June			July		
Topics	PMI	NPMI	TD	PMI	NPMI	TD
10	1.705	-0.085	0.66	1.646	-0.084	0.64
20	2.196	-0.112	0.55	1.944	-0.099	0.54
30	2.425	-0.125	0.46	2.559	-0.132	0.46
50	2.94	-0.154	0.46	3.021	-0.159	0.42
70	3.011	-0.159	0.40	3.299	-0.174	0.45
90	3.357	-0.179	0.28	3.437	-0.182	0.39
110	3.58	-0.192	0.33	3.561	-0.19	0.38

optimally in subtopics.

The purpose of the study is to identify the most popular and diverse topics of discussion, thus the tweets in the corpus were grouped in 10 topics for each month and the composition of these topics in terms of keywords and their distribution within the corpus was further analyzed. The choice to use 10 is also motivated by the observation of Marinov et al. (2020) which indicated the value of NPMI is a good discriminator of an optimal number of topics. Therefore, the choice of the topics number maximized the NMPI value.

3.2 Topic Modelling Results

The SeaNMF algorithm was used to identify 10 topics each month between February and July. The topics were named in accordance with the list of their top keywords - most frequent words occurring in each word cluster. Thus, when similar lists of top 10 keywords occurred for topics identified in different months, the topics were considered to be the same topic, even though their less popular keywords might differ. For instance, the top 10 keywords for the *Arts* topic in April are *arte, artista, raccontare, autore, protagonista, museo, poesia, musico, bellezza, opera* (Eng. arts, artist, to tell/narrate, author, protagonist, museum, poem, music, beauty, opera), while in March they are artista, arte, celebrare, raccontare, antico, capolavoro, protagonista, club, autore, streaming (Eng. artist, arts, celebrate, to tell/narrate, ancient, masterpiece, protagonist, club, author, streaming). The two word clusters were considered to define the same topic, since they share several top keywords (*artista, arte, raccontare, autore, protagonista*) and the remaining words are semantically related (*capolavoro* and *opera*, since *opera* refers both to a musical genre and a masterpiece).

3.3 Topics Distribution

Figure 3.1 illustrates the distribution of the popular topics identified by the SeaNMF model each month. The X-axis indicates the dates and the Y-axis indicates the number of tweets classified as concerning a particular topic, on a particular date. The graph indicates that there is a much more popular topic than the others. In fact, the *Covid-19 cases* topic occurs in at least 30% of the tweets each month, and in over 50% in February and April. The distribution of the tweets concerning *Covid-19 cases* follows quite closely the overall distribution of the tweets in the corpus, as illustrated by

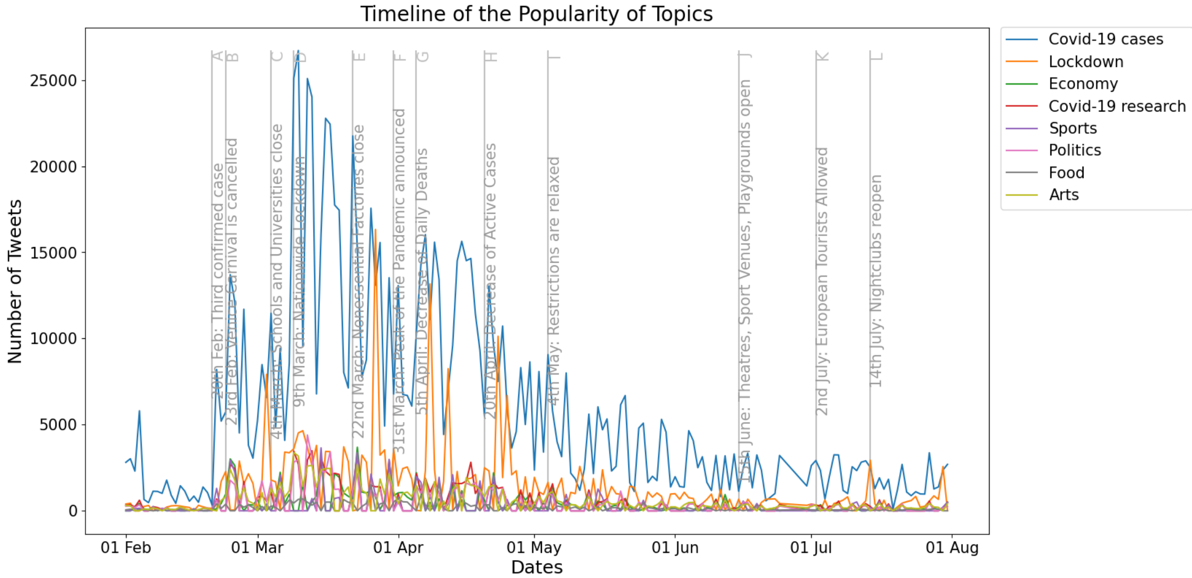

Figure 3.1: Timeline of the distribution of the most popular topics each month.

Figure 2.1. Thus, there are spikes in the popularity of *Covid-19 cases* discussions after the Venice Carnival was cancelled (February 23rd), after the nation-wide lockdown was issued (March 9th),and after factories were closed (March 22nd). The next most popular topic is the *Lockdown* topic, which first gained popularity before schools were closed on March 4th and then spiked irregularly, at times when the *Covid-19 cases* was also spiking, but was never more popular than the discussion about the *Covid-19 cases*. Quite surprisingly, even after the lockdown was issued, the *Lockdown* topic itself did not gain more popularity. Similarly, after theaters, sports and playgrounds opened, there was no spike in discussions about *Sports* or *Arts*; however, when the nightclubs re-opened there was a spike in discussions about the *Lockdown*.

The 5 most popular topics between February and April are also shown in Table 3.2, alongside the popular topics identified by Marinov et al. (2020) in their study of Dutch tweets. The comparison between the findings of the studies is discussed in Section 4.3.

3.4 Emotion Detection

The emotion scores conveyed by each tweet were computed on a look-up basis, using the NRC EmoLex (Mohammad and Turney, 2013) for words, and the EmoTag lexicon (Shoeb et al., 2019) for emojis, and summing the scores of individual words and emojis in a tweet. A score of 0 for a particular emotion means that the tweet does not convey that emotion, while a score equal to 1 means that the emotion is clearly expressed in that tweet.

Not all tweets conveyed an emotion, according to the lexicons. However, there were 1 830 089 tweets with words that conveyed emotions, 191 087 tweets with emojis that conveyed emotions, and 134 329 tweets which contained both words and emojis that conveyed emotions. Since emotions were only identified in about 60% of tweets, it was not possible to continuously plot the emotions associated with each topic - except in the case of the most popular topic illustrated in Figure 3.2 and analysed in the next paragraph. Thus, the average emotion scores per month for each topic were illustrated in tables in Appendix A. The tables suggest that regardless of topic, there is a similar pattern in the intensities of the emotions conveyed. In fact, fear seems to be the most accentuated emotion, followed by trust, anticipation, sadness, then joy, anger and surprise with quite similar intensities, and finally disgust. This could hint towards a bias of the words included in the lexicons, which will be later analysed in the Discussion section.

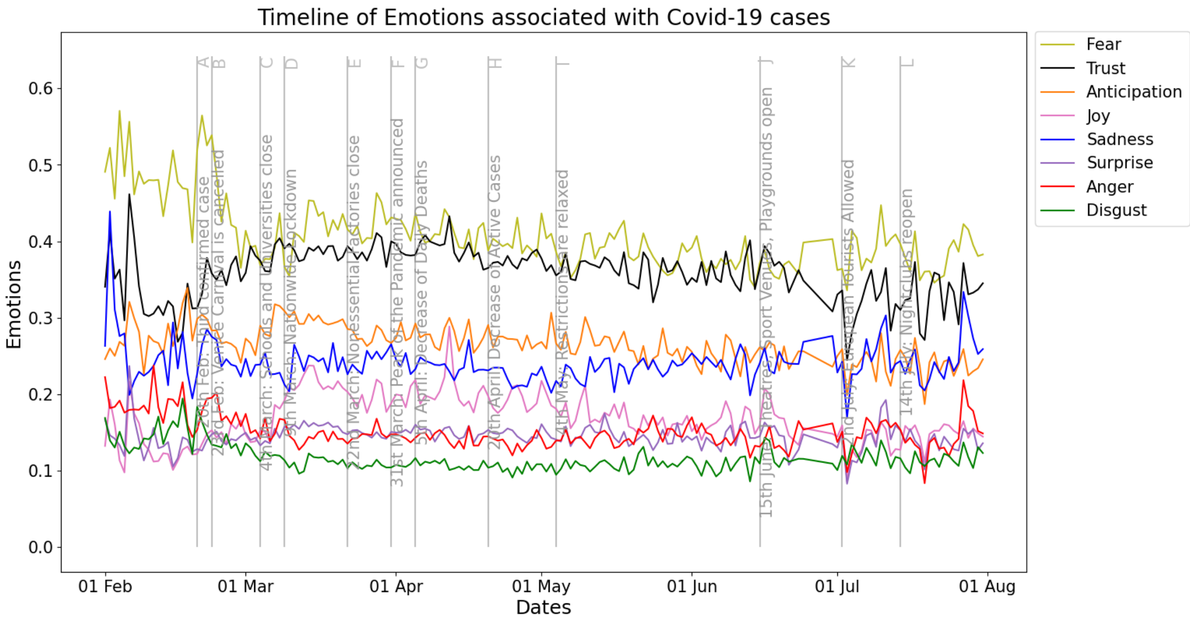

Figure 3.2: Timeline of emotions associated with the *Covid-19 cases* topic. The emotion scores are computed as the sums of the emotion scores identified using the NRC EmoLex and the EmoTag lexicon, averaged per day.

Figure 3.2 illustrates the emotion ratings of the tweets concerning the most frequent topic between February and July 2020. The X-axis shows the dates and the Y-axis shows the average emotion score per day for a particular emotion, computed using both lexicons. It seems that the intensity of the emotions related to the *cases of Covid-19* is rather constant, especially during the lockdown period, between March 9th and June 15th. Tweets from this period convey an accentuated sense of fear, complemented by trust, which scores a slightly lower emotional score. Joy seems to be more present in tweets before restrictions were relaxed on May 4th. Overall, disgust is the least intensely conveyed emotion. There is a noticeable spike in fear intensity just after February 20th, when the third case of infection with Covid-19 was confirmed in Italy. Furthermore, there is a noticeable dip in the intensity of all emotions just after July 2nd, when European tourists were allowed to travel to Italy again. However, this is not a meaningful pattern, since on this date there was also a drop in the number of tweets concerning the topic of *Covid-19 cases*, as suggested by Figure 3.1.

3.5 Emotion Polarity

The polarity score indicates the directionality of the emotions conveyed by a tweet. It is computed by summing the scores of positive emotions (trust, joy), and then subtracting the scores of negative emotions (fear, anger, disgust, sadness). Thus, a negative polarity score indicates indicates a mostly negative attitude, while a positive score indicates a mostly positive one, and polarity equal to 0 suggests a neutral attitude.

Figure 3.3 illustrates the polarity of topics, according to the EmoTag lexicon, which classifies emojis. Neither of the graphs can provide a continuous value for the polarity of each topic, since not all topics occurred daily in tweets. According to the EmoTag lexicon, the polarity of most topics was close to neutral, oscillating between small positive and negative values. The lockdown topic seems to have yielded some of the most polarized tweets. The positive peak which occurred in the beginning of April is likely linked to the announcements about the decrease in daily infections and deaths. The accentuated negative polarity which occurred in the tweets about the lockdown during the mid-

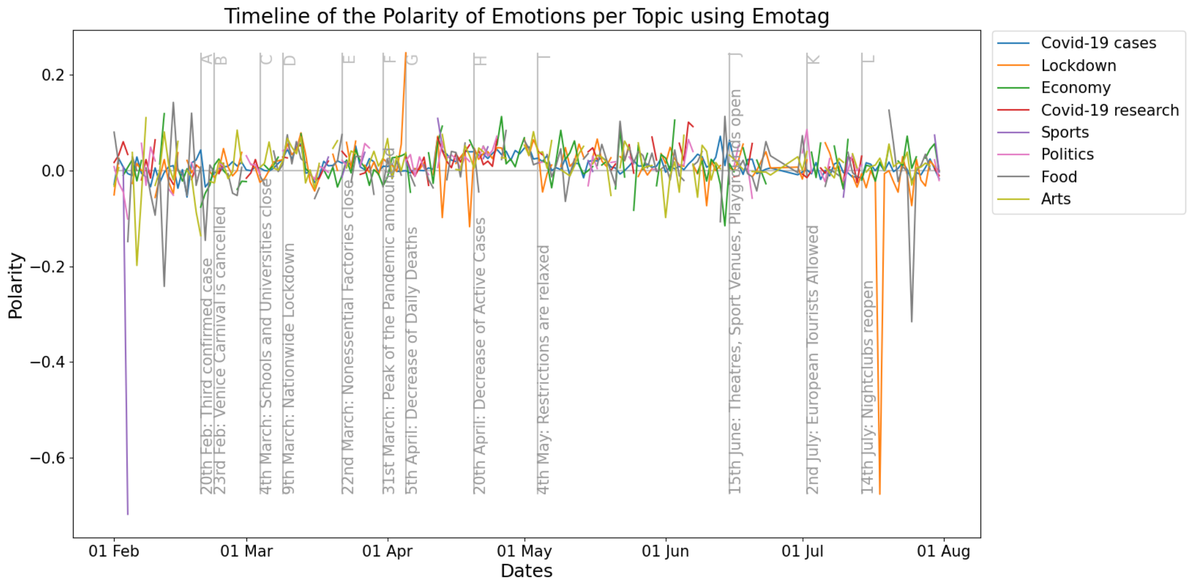

Figure 3.3: Polarity of the emotions identified using the EmoTag lexicon.

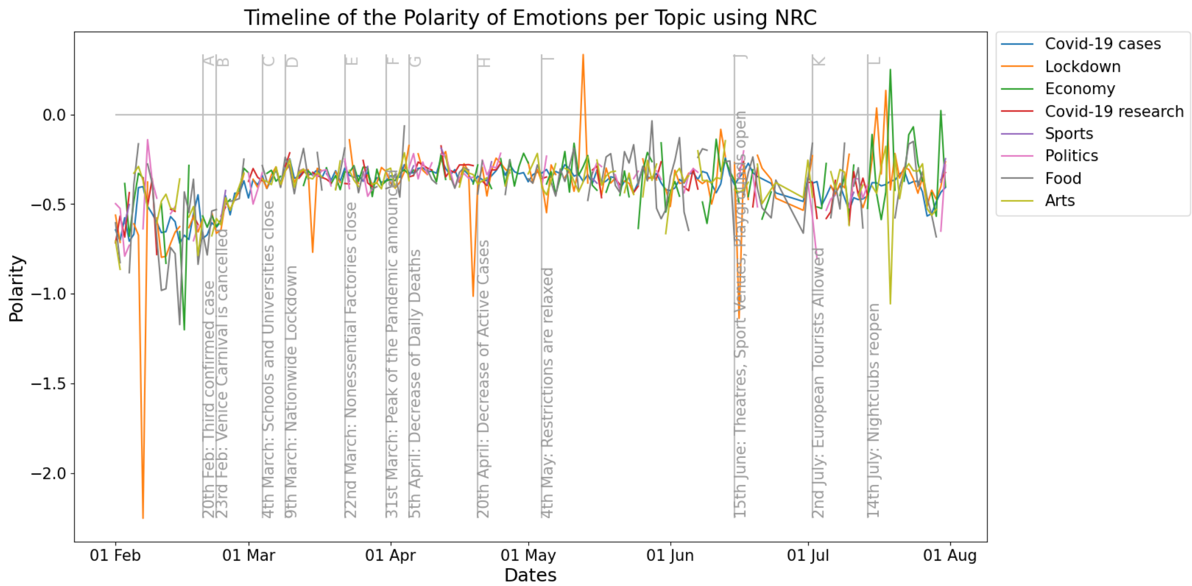

Figure 3.4: Polarity of the emotions identified using the NRC EmoLex.

dle of July is likely linked to the announcements about loosening the restrictions and re-opening of the night-clubs.Therefore, the emojis used in tweets suggest that Italians were largely content with the lockdown and favored stricter restrictions that ensured the safety of the general population.

Another topic which registered large oscillations in emotion polarity was the topic related to food. News reports (Zhu, 2020) note that panic buying and food shortages started as early as February 20th, following the confirmation of the first infection with Covid-19 in Italy. The most negative polarity was observed in early February, in connection to the sports topic and it is most likely related to the football matches of the Serie A - an Italian national football league - which took place in that period, before being suspended at the start of the lockdown.

Figure 3.4 illustrates the polarity of various topics, computed according to the NRC EmoLex, which only accounts for the emotions conveyed by words. Overall, the words used in tweets conveyed more negative emotions than the emojis did. Similar to the case of emojis, the lockdown topic registered the most ample oscillations. The dip in polarity registered in February is likely due to discussions about early measures to prevent the spread of the virus, as the first cases were confirmed in Lombardy. However, the same pattern was not observed in the polarity of tweets. In fact, most of the dips do not coincide, which suggests that the emotions conveyed through emojis are not necessarily correlated to the ones expressed through words. There is a notable peak that indicates positive attitudes in tweets about the state of the economy in the middle of July, which is likely related to the announcement that the European Commission approved a financial aid to support various Italian economy sectors (European Comission, 2020).

4 Discussion

This study aimed to examine the topics and emotions present in Italian tweets collected between February and July 2020, in order to assess how the discourse surrounding the Covid-19 pandemic was influenced by notable events which occurred in that time-frame. To this end, a SeaNMF model was trained to identify the major topics of discussion on Twitter and the tweets in the corpus were classified accordingly. The findings suggest that the distribution of topics was imbalanced, and found little evidence for a direct connection between changes in topic trends and events. These findings are further discussed in Section 4.1.

Additionally, the emotional rating of each tweet was computed according to the words and emojis it contained, in order to compare the emotional undertones of the identified topics. The results in this case surprisingly indicated a similar distribution of emotions for all of the major topics, fear and trust being the most emphasised emotions, regardless of the context in which they were expressed. Moreover, for the same topics, the polarity of emotions conveyed through words was found to be more negative than the polarity of emotions conveyed through emojis. These findings may hint that the resources used for emotion detection were not reliable, as discussed in Section 4.2.

Next, a comparison is made between the topic modelling findings and the results reported by Marinov et al. (2020), who conducted a similar analysis on Dutch tweets collected between February and May (Section 4.3). Finally, the possible directions for future research and explored in Section 4.4, while improvements to the methodology of the current study are presented in Section 4.5.

4.1 Topic Modelling Findings

Given that the collection of the tweets in the 40wita corpus was based on a series of keywords related to the Covid-19 pandemic, some of the topics which were identified in the tweets were directly related to the pandemic (such as *Covid-19 Cases*, *Covid-19 Research*, *Lockdown*). However, Italians also discussed topics such as *Politics* and *Economy* in the context of the pandemic, and even associated *Sports* and *Arts* with the keywords used by the 40wita corpus.

The most popular topic (among the topics that occurred monthly) was concerned with the newly reported Covid-19 infections. This suggests that a considerable number of tweets each month were content-wise similar to news reports about the development of the pandemic. Many of these tweets were likely posted by media outlets or government agencies with the intent of keeping the general population up to date with the latest information about the pandemic. This phenomenon is in line with the findings of Daughton and Paul (2019) and Househ (2016), which indicated that Twitter had also been an efficient means of communicating important news during both the Zika and Ebola outbreaks.

Given that the most popular topic surrounding the Covid-19 cases accounts for between 30% and 50% of tweets each month, it largely follows the distribution of the corpus, registering more tweets during the days when more tweets were collected for the corpus. Thus, although there is a tendency

Table 3.2: Table of the most popular topics between February and April, compared to the most popular topics found by Marinov et al. (2020)

February		March		April	
Italy	Netherlands	Italy	Netherlands	Italy	Netherlands
Covid-19 Cases	Covid-19 China	Covid-19 Italy	Covid-19 Netherlands	Covid-19 Cases	Economy
Lockdown	Early Covid-19	Covid-19 Cases	Infections	Lockdown	Government
Covid-19 Research	Global Issues	Lockdown	Economy	Economy	Global Issues
Covid-19 Testing	Covid-19 Europe	Covid-19 Research	Government	Politics	Measures
Sports	Measures	Economy	Global Issues	Arts	Covid-19 Europe

for the number of tweets concerning Covid-19 cases to increase on days when notable events related to the pandemic occurred, there are also many dates on which the number of tweets peaks, but no important event was registered. Furthermore, no statistical tests were used to examine a possible correlation, therefore there is no evidence to support that the popularity of the topics within the discourse surrounding the Covid-19 pandemic was influence by notable events.

By examining how the keywords of a topic change from month to month, it can be inferred how topics shifted their focus over time. For instance, the top 5 keywords of the *Covid-19 cases* in February were *caso, italia, primo, contagiare, cina*, which reveal that at that point the coronavirus was still strongly associated with China (*cina*) and that discussions were mostly concerning the first cases (*primo*) in Italy. However, as the infections spread in March, the popular keywords became *caso, positivo, nuovo, oggi, guarire*, which indicate a focus on new cases (*nuovo, caso*) and on the survivability rate (*guarire* means to heal/recover). At this point, the infections were no longer associated with China as much and nor was *Italia* a top keyword, since the magnitude of infections in Italy was already established. In the following months, the top keywords remained constant (*nuovo, caso, positivo, decesso, morto*), revealing a focus on the number of new infections (*nuovo, caso, positivo*) and on the number of deaths due to Covid-19 (*decesso, morto*).

The keywords used to discuss the lockdown have changed drastically. In February, the keywords were *regionale, misura, urgere, disposizione, ordinanza*, as the topic was only covering regional measures (*regionale, misura*) adopted by several municipalities in Northern Italy. Furthermore, the terms used to describe the measures are conveying urgency, severity and authority (*urgere, disposizione, ordinanza*). In March, once the lockdown was extended to the entire country, the focus of the topic shifted towards reinforcing that people should stay home and follow the restrictions through words such as *casa, giorno, restare, stare, orare*, meaning: "home", "day", "to remain (at home)", "to stay (at home)", "to pray". Finally, in June, as Italy entered Phase Three of the lockdown and the most severe restrictions were cancelled, the top keywords of the discussions surrounding the lockdown were *pandemia, anno, causare, lockdown, stare* ("pandemic", "year", "to cause", "to stay"), which suggest that the conversations were looking back, analysing the causes of the pandemic.

4.2 Emotion Detection Findings

Similar patterns in the emotion ratings were registered for all topics investigated, as seen in the tables from Annex A. Most notably, fear and trust scored the highest average values, however, the level of fear decreased gradually between February and July while the level of trust peaked either in March or in April for all topics. Given that the lockdown was imposed on the 9th of March, and Phase Two - when the restrictions were relaxed - began only on the 4th of May, the period in which the level of trust was the highest was during the stricter lockdown. This might suggest that Italians generally felt safer during the lockdown, as their tweets conveyed slightly more trust.

It is important to note that the emotion scores are largely based on the ratings of the NRC EmoLex, given that about 96% of the emotion scores were computed using the NRC EmoLex alone, as relatively few tweets contained emojis. However, the NRC EmoLex was composed for the English language and translated to Italian (Mohammad and Turney, 2013), which might have lead to unreliable emotion scores and bias. Furthermore, the NRC EmoLex does not differentiate between word sense or parts of speech, which limited the possibilities for disambiguation. Therefore, the emotion detection findings may not be an accurate representation of the emotions conveyed by the tweets.

There was a noticeable difference between the polarity of the emotions detected using the NRC EmoLex and the EmoTag lexicons, with the former yielding a negative polarity and the latter a more neutral polarity. This difference might be arguably accounted for by the unreliability of the NRC EmoLex for foreign languages, or it might point towards an inherent difference between the sentiments and the types of emotions that are conveyed through words and emojis.

4.3 Comparison with Dutch Tweets

Marinov et al. (2020) performed topic modelling on a set of Dutch tweets collected between February and April 2020, examining different the topics discussed by different demographic groups on Twitter, and their emotions towards the topics. They also reported the most popular topics among all categories of users, which are shown in Table 3.2, along with the most popular Italian topics during that period.

In terms of the variation in the most popular discussion themes between February and April, Italian tweets were focused on a set of 8 unique topics, while the Dutch tweets covered a range of 9 unique topics. Italian tweets were focused on reporting the daily numbers of infections in Italy (*Covid-19 Cases*), and discussing the restrictions imposed by the government (*Lockdown* and to some extent *Covid-19* Italy). These topics also come up in Dutch tweets as *Infections*, *Measures*, and to some extent covered by *Covid-19 Netherlands.* The state of the economy was also discussed by both Twitter communities, although the topic was slightly more popular in the Netherlands. Politics and the government are also similar topics which were addressed by both Dutch and Italian Twitter users.

Dutch tweets were much more preoccupied with the development of the pandemic globally (*Covid-19 China*, *Covid-19 Europe*), and with *Global Issues*, while in Italy the discussion surrounding the pandemic was mostly focused on internal affairs (*Covid-19 Cases*, *Covid-19 Testing*, *Lockdown*). Furthermore, Italians consistently discussed the advancements in Covid-19 research, but the Dutch tweets did not address this topic. Given that in Italy the Covid-19 infections were more widespread and the lockdown was stricter than in the Netherlands, it is understandable that the Italian tweets were more focused on internal affairs than Dutch ones and that they were preoccupied with research into the virus.

Lastly, the popular Italian topics also feature subjects related to entertainment and hobbies such as *Sports* and *Arts*. The topic of *Sports* was popular in February due to the Serie A football championship, which was postponed due to the rise in infections. However, there was no particular event related to *Arts* and the popularity of this topic might suggest that Italians spend more time enjoying their hobbies during the lockdown.

4.4 Future Research

As previously mentioned, the most popular topic found on the Italian Twitter during the study was regarding the spread of Covid-19 infections, which might suggest that some of these tweets originated from the Twitter accounts of media outlets or health authorities. Therefore, the emotions and attitudes conveyed by these tweets might not reflect the opinions of the general population. A natural next step would be to classify the tweets of the 40wita corpus based on the characteristics of the users who posted them. Thus, the topics and emotions of the general population can be examined independently of the official reports.

In order to assess the links between notable events and trends in topic popularity and emotions, more dates can be considered as reference points. When the results were analysed, some of the emerging trends were traced back to events which were not included in the initial set of notable events (Table 2.1). Therefore, a more efficient workflow would be to identify important dates based on the observed trends. Furthermore, focusing on tweets posted within a short time-frame around a notable event could lead to a better understanding of the impact of the event on the users' discourse.

4.5 Improvements

As explained in the Results section, the evaluation metrics of the topic model indicate that the more topics are used, the less diverse they become (as suggested by the TD score) and the less likely their keywords are to co-occur in a tweet (as suggested by the NPMI score).

A possible solution to achieve a better trade-off between the coherence and diversity of the topics could be changing the criteria for choosing keywords. The method used, which was also employed by Marinov et al. (2020) and Shi et al. (2018), was based on the frequency in the corpus. However, in a situation when one topic is much more frequent than the others, many of its keywords will also be very frequent which will result in a set of keywords that are not representative for most of the topics in the corpus. The skewed choice of keywords will also negatively impact evaluation metrics. Since most keywords belong to a single topic, the topic diversity decreases when more topics are considered. The topic coherence measures are also more likely to indicate a low correlation or even an independence between the keywords within the same topic, since all keywords have a high frequency and are more likely to occur in various contexts, than in fixed patterns.

Therefore, choosing the keyword features of a SeaNMF model should not be based on the word frequency alone, as an imbalanced distribution of topics could impair the ability of the model to identify more diverse and coherent topics. Instead, the term frequency–inverse document frequency (TF-IDF) score can be used, which will highlight slightly less frequent, but more informative words and penalize the very frequent and less informative words. This approach has been used by LDA and NMF models, although it can distort the semantic coherence of topics (Suh, Choo, Lee, and Reddy, 2016). However, in the case of the SeaNMF algorithm - which is specifically designed to achieve semantically sound results by exploiting word contexts - the possible disadvantage of using the TF-IDF criteria could be negligible.

Nevertheless, in the case that the findings are indeed an accurate representation of the distribution of topics in the 40wita corpus, it is also possible that the keywords used to select the tweets included in the corpus are biased towards certain topics. Thus, the 40wita corpus may not fully capture the diversity of topics discussed by Italians during the early Covid-19 pandemic.

4.6 Conclusions

The study attempted to examine how the topics of discussion changed in Italian tweets between February and July 2020. The findings suggest that the topics which were consistently popular during that period were related to the Lockdown and the spread of Covid-19 infections. Furthermore, an analysis of the keywords associated with these topics reveals that the focus of the topics did change between February and July. Although these changes might be correlated with different stages of the pandemic and the lockdown, there is not enough evidence to support such claims. In addition, the comparison between popular topics in Italian and in Dutch tweets from the same time-frame indicates that Italians were more preoccupied with internal affairs (most notably the development of the pandemic on a national scale) than the Dutch.

The analysis of the emotions related to each topic suggests that all topics had a similar distribution of the emotions associated with them, with fear and trust being the most emphasized and disgust being the least. There is a trend across all topics that points towards the levels of trust and fear reaching similar values during the (stricter portion of the) lockdown. Finally, the polarities yielded by the two lexicons used for the emotion detection task suggest that the emotions assigned by the NRC EmoLex are more negative than the ones assigned by the EmoTag lexicon; such discrepancies raise questions about the reliability of the resources used for the study.

References

Valerio Basile and Tommaso Caselli. 40twita 1.0: An collection of italian tweets during the covid-19 pandemic, 2020. URL `http://twita.di.unito.it/dataset/40wita`.

Valerio Basile, Mirko Lai, and Manuela Sanguinetti. Long-term social media data collection at the university of turin. In *Proceedings of the Fifth Italian Conference on Computational Linguistics (CLiC-it 2018), Torino, Italy, December*

10-12, 2018., 2018. URL http://ceur-ws.org/Vol-2253/paper48.pdf.

Steven Bird, Ewan Klein, and Edward Loper. *Natural language processing with Python: analyzing text with the natural language toolkit.* " O'Reilly Media, Inc.", 2009.

Megan R. Brett. Topic modeling: A basic introduction. *Journal of Digital Humanities*, 2(1), 2012.

Jaegul Choo, Changhyun Lee, Chandan Reddy, and Haesun Park. Weakly supervised nonnegative matrix factorization for user-driven clustering. *Data Mining and Knowledge Discovery*, 29, 11 2015. doi: 10.1007/s10618-014-0384-8.

Ashlynn R Daughton and Michael J Paul. Identifying protective health behaviors on twitter: Observational study of travel advisories and zika virus. *J Med Internet Res*, 21(5):e13090, May 2019. ISSN 1438-8871. doi: 10.2196/13090. URL https://www.jmir.org/2019/5/e13090/.

Ensheng Dong, Hongru Du, and Lauren Gardner. An interactive web-based dashboard to track covid-19 in real time. *Lancet Infectious Diseases*, 20:533–534, 2020. doi: doi:10.1016/S1473-3099(20)30120-1.

European Comission. State aid: Commission approves €6.2 billion italian grants scheme to support small businesses and self-employed affected by coronavirus outbreak [press release], July 2020. URL https://ec.europa.eu/commission/presscorner/detail/en/IP_20_1311.

Matthew Honnibal and Ines Montani. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. 2017.

Mowafa Househ. Communicating ebola through social media and electronic news media outlets: A cross-sectional study. *Health informatics journal*, 22:470–478, 2016. doi: https://doi.org/10.1177/1460458214568037.

Daniel D. Lee and H. Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.

Giovanni Legorano and Eric Sylvers. As virus spreads, italy locks down country. *Wall Street Journal*, Mar 2020. URL https://www.wsj.com/articles/italy-bolsters-quarantine-checks-after-initial-lockdown-confusion-11583756737.

Boris Marinov, Jennifer Spenader, and Tommaso Caselli. Topic and emotion development among Dutch COVID-19 Twitter communities in the early pandemic. In *Proceedings of the Third Workshop on Computational Modeling of People's Opinions, Personality, and Emotion's in Social Media*, pages 87–98, Barcelona, Spain (Online), December 2020. Association for Computational Linguistics. URL https://www.aclweb.org/anthology/2020.peoples-1.9.

Cameron Martel, Gordon Pennycook, and David G. Rand. Reliance on emotion promotes belief in fake news. *Cognitive research: principles and implications*, 5, 2019. doi: https://doi.org/10.31234/osf.io/a2ydw.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space, 2013a.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 26, pages 3111–3119. Curran Associates, Inc., 2013b.

Saif M. Mohammad and Peter D. Turney. Crowdsourcing a word-emotion association lexicon. 29 (3):436–465, 2013.

Nic Newman, Richard Fletcher, Anne Schulz, Simge Andı, and Rasmus Kleis Nielsen. Reuters institute digital news report 2020, 2020.

Catherine Ordun, Sanjay Purushotham, and Edward Raff. Exploratory analysis of covid-19 tweets using topic modeling, umap, and digraphs. 2020.

Robert Plutchik. A general psychoevolutionary theory of emotion. In *Emotion: Theory, research* *and experience, Theories of emotion*, volume 1, pages 3–33. Academic Press, 1980.

Tian Shi, Kyeongpil Kang, Jaegul Choo, and Chandan Reddy. Short-text topic modeling via nonnegative matrix factorization enriched with local word-context correlations. pages 1105–1114, 04 2018. doi: 10.1145/3178876.3186009.

Abu Awal Md Shoeb, Shahab Raji, and Gerard de Melo. EmoTag – Towards an emotion-based analysis of emojis. In *Proceedings of RANLP 2019*, pages 1094–1103, sep 2019. doi: 10.26615/978-954-452-056-4_126. URL https://www.aclweb.org/anthology/R19-1126.

S. Suh, J. Choo, J. Lee, and C. K. Reddy. L-ensnmf: Boosted local topic discovery via ensemble of nonnegative matrix factorization. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*, pages 479–488, 2016. doi: 10.1109/ICDM.2016.0059.

Jia Xue, Junxiang Chen, Ran Hu, Chen Chen, Chengda Zheng, Yue Su, and Tingshao Zhu. Twitter discussions and emotions about covid-19 pandemic: a machine learning approach. *Journal of medical Internet research*, 2020. doi: 10.2196/20550.

G. Xun, V. Gopalakrishnan, Fenglong Ma, Y. Li, J. Gao, and A. Zhang. Topic discovery for short texts using word embeddings. *2016 IEEE 16th International Conference on Data Mining (ICDM)*, pages 1299–1304, 2016.

Yuru (Priscilla) Zhu. February 2020. URL https://www.businessinsider.nl/italy-coronavirus-surge-stock-up-supplies-supermarkets-photos-2020-2?international=true&r=US.

A Tables of Average Emotion Scores per Month, for various Topics

Table A.1: Table for *Covid-19 Research* topic

	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
February	0.182609	0.261575	0.137482	0.454456	0.140489	0.253254	0.131887	0.353417
March	0.145339	0.287068	0.110256	0.410962	0.197447	0.241325	0.152497	0.385862
April	0.138825	0.274567	0.104636	0.400505	0.195554	0.230499	0.151278	0.39194
May	0.138395	0.252385	0.104027	0.384957	0.174328	0.226336	0.140458	0.36062
June	0.13373	0.253394	0.111956	0.369181	0.171229	0.2416	0.153273	0.361146
July	0.150603	0.239468	0.115506	0.381814	0.139716	0.246893	0.141157	0.327727

Table A.2: Table for *Covid-19 cases* topic

	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
February	0.181224	0.275004	0.140034	0.478441	0.14527	0.259697	0.140969	0.353368
March	0.146557	0.289245	0.110607	0.409749	0.198484	0.238633	0.152856	0.386137
April	0.141428	0.272954	0.106402	0.405473	0.196612	0.234793	0.148545	0.386542
May	0.143217	0.261624	0.107221	0.38714	0.172016	0.22911	0.144637	0.359047
June	0.143558	0.252511	0.112294	0.372227	0.165008	0.24439	0.143953	0.360147
July	0.151532	0.242616	0.117525	0.384022	0.143465	0.249413	0.141878	0.327775

Table A.3: Table for *Sports* topic

	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
February	0.174305	0.270303	0.130853	0.464611	0.139054	0.250182	0.135678	0.345348
March	0.143834	0.272314	0.111827	0.416148	0.192696	0.242336	0.146321	0.376745
April	0.137437	0.276009	0.110073	0.411253	0.205614	0.233967	0.152378	0.391499
May	0.145376	0.280917	0.11217	0.404923	0.184848	0.238033	0.152927	0.372429
June	0.146135	0.25499	0.105593	0.362059	0.181349	0.23475	0.15479	0.374293
July	0.145373	0.250003	0.106193	0.37332	0.159989	0.242754	0.160953	0.351844

Table A.4: Table for *Politics* topic

	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
February	0.194066	0.270069	0.145599	0.468037	0.141036	0.258771	0.139016	0.352365
March	0.144427	0.287433	0.109357	0.404478	0.198742	0.234412	0.152902	0.379195
April	0.138974	0.278686	0.106745	0.401635	0.201498	0.234807	0.150962	0.390161
May	0.143893	0.270882	0.103191	0.391952	0.180975	0.230652	0.14562	0.381968
June	0.137401	0.252411	0.105523	0.365628	0.171398	0.231233	0.143662	0.349178
July	0.159561	0.243024	0.116042	0.378857	0.152934	0.252487	0.143098	0.323464

Table A.5: Table for *Economy* topic

	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
February	0.177508	0.262886	0.139852	0.454549	0.144792	0.255488	0.136123	0.351564
March	0.144278	0.285641	0.111883	0.420016	0.195282	0.245205	0.150511	0.387546
April	0.1429	0.277095	0.106041	0.418837	0.208236	0.234674	0.151828	0.394539
May	0.146819	0.27614	0.110004	0.393878	0.169855	0.231514	0.147866	0.370501
June	0.141608	0.246297	0.104947	0.390888	0.151886	0.251983	0.148736	0.352785
July	0.139975	0.237535	0.109697	0.369988	0.159687	0.245936	0.150848	0.317508

Table A.6: Table for *Lockdown* topic

	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
February	0.181404	0.269448	0.14111	0.46562	0.14362	0.253449	0.141291	0.357165
March	0.146762	0.284293	0.110397	0.409433	0.192239	0.240705	0.148385	0.386861
April	0.137282	0.274262	0.103773	0.407525	0.204161	0.228944	0.150978	0.391896
May	0.141507	0.266466	0.103499	0.392072	0.177274	0.229184	0.143903	0.364331
June	0.144833	0.253993	0.111589	0.371343	0.162573	0.242596	0.146533	0.360443
July	0.148572	0.244528	0.115918	0.372261	0.148491	0.247372	0.145102	0.33224

Table A.7: Table for *Food* topic

	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
February	0.193671	0.278869	0.157933	0.479044	0.143482	0.266436	0.149826	0.34645
March	0.14492	0.296526	0.111284	0.409701	0.193057	0.243087	0.146014	0.383809
April	0.150815	0.277458	0.109963	0.42154	0.203436	0.247823	0.151899	0.389216
May	0.137577	0.269015	0.1135	0.379981	0.17292	0.236502	0.148313	0.360504
June	0.138287	0.247145	0.103009	0.376694	0.141845	0.235269	0.132476	0.347141
July	0.151389	0.277308	0.113984	0.396413	0.168004	0.266356	0.152444	0.358179

Table A.8: Table for *Arts* topic

	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
February	0.179076	0.272336	0.137085	0.463753	0.144814	0.255019	0.142416	0.356001
March	0.146759	0.293894	0.110632	0.415913	0.205426	0.238946	0.154777	0.390039
April	0.140626	0.269779	0.105421	0.405653	0.196098	0.233415	0.150078	0.380724
May	0.144129	0.265109	0.105687	0.395813	0.181235	0.234819	0.147319	0.37023
June	0.14902	0.271788	0.112514	0.38822	0.17383	0.256569	0.162408	0.379682
July	0.152688	0.246081	0.113386	0.373842	0.153553	0.24357	0.146153	0.335888