Matriculation number: 200007413

How has (or could) automatic text analysis been used in the handling of the COVID-19 crisis?

Matriculation Number: 200007413
Tutor: Oggie Arandjelovic
Word count: 1989

Matriculation number: 200007413

Introduction

COVID-19 has affected our lives in unprecedented ways, all in such a small period of time. With this pandemic, it has presented the first instance of a global pandemic alongside widespread social media usage. During the last pandemic (the 2003 SARS virus), the main form of digital communication was e-mail (Griffith et al. 2003). Now, we have had new tools for spreading information and gauging the public response to new measures. Yet so too have novel problems arisen, such as the need to tackle misinformation - which negatively impacts public compliance and exacerbates incivility. This is why social media corporations have attempted to tackle the issue of misinformation on their platforms and social analysis attempted to gather information regarding public trust & perceptions through the news and social media. Plus, why governments potentially should have utilized these analysis tools in guiding decisions on tackling the virus, whilst keeping people on-side by addressing the main concerns of the populace.

Automatic Text Analysis – An Overview & Issues with COVID-19

The aim of automatic text analysis is to semantically categorize and understand the content of pieces of text. As language tends to provide reasonably structured, with constraining rules, and clearly semantic words or phrases: It is possible for a machine to mathematically gain insights into the meaning of pieces of text.

There are various methods of automatic text analysis which have different potential uses for tackling different problems. Many of these have been used in tackling different issues COVID-19 has presented over time which I will discuss in more detail later. I will briefly describe the different automatic text analysis techniques which are commonly used and have been implemented in different COVID-19 related example. Firstly, semantic/emotional text analysis is where the general positive,negative or neutral feeling a piece of text appears to be presenting can be mathematically determined by analyzing the semantics of words and phrases with known strength of emotion. Further, emotional text analysis can determine finer detail emotions a piece of text presents through analysis of the emotions the words and phrases in a piece of text usually are presenting using mathematical analysis. The last example of automatic text analysis is feature analysis: where text can be broken down to key features, most commonly using a 'bag of words' which uses a dictionary of words which are or are not present in a piece of text which can be compared to other pieces of text. This can be applied by having a database of known topic pieces of text and new text can be compared to mark related content.

As the examples that will be discussed in this essay are focused on posts to social media and news sites, a vital step with every form of automatic text analysis is pre-processing. Whilst there are various methods of pre-processing (stemming, lemmanization, stop-word removal, etc.), the details of such methods are unrelated to the topic of this essay. However, when it comes to these platforms and the current crisis, the importance of these steps are crucial as to allow the analysis to be efficient and scalable to process every post. Even tiny increases to efficiency, scaled over the entirety of a platform will have a huge impact. The novel nature of the pandemic means that the content of posts is ever changing and new issues are constantly arising. Being able to keep up with the current posts allows the tools to remain practically useful.

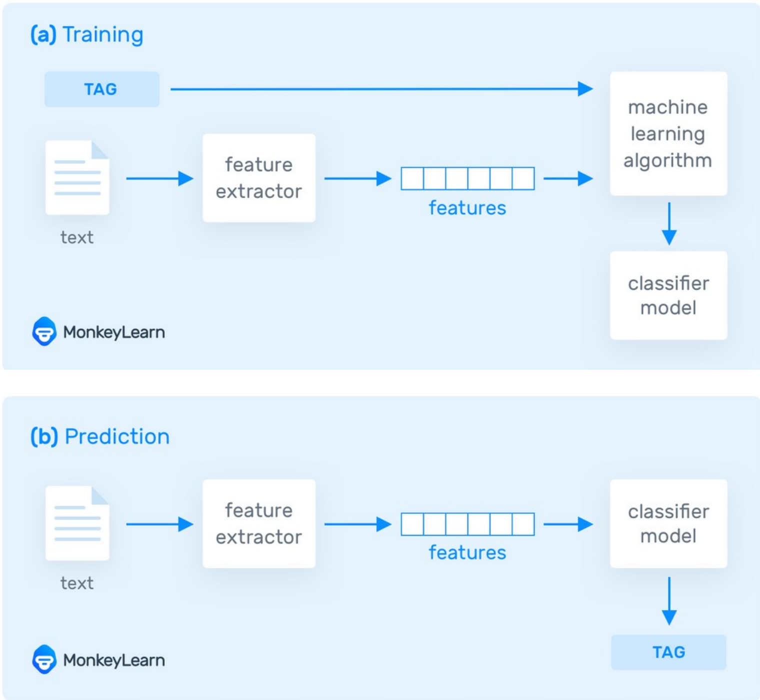

Figure 1: How automatic feature analysis can be implemented in automatic tagging of text. This can be useful for search engines for detecting the topic of content.

However, COVID-19 presents a difficulty with implementing these tools. As much of the time a dictionary of important words and phrases are needed for feature extraction to work, and with new terminology, slang and phrases developing as the crisis continues, it is difficult to automatically add these new important features to the dictionary. If the tool is supposed to be completely automated then it may miss these changes in the conversation around the pandemic and only focus on the topics that we are already aware of. This is one factor in potentially why governments have not decided to officially take on board the tools that have been developed over the time of the crisis as it could bring about accusations of bias if used to determine what issues to focus on.

Tackling Misinformation – Information Spread on Social Media

Misinformation is a well-known, and seemingly ever growing issue with social media - in regards to notoriety and impact (Wang et al. 2019). In particular, fact-checkers found an extremely sharp increase in fact-checks relating to COVID-19 in the first few months of 2020, some by over 900% (Brennen et al. 2020). Clearly, social media companies needed a response to prevent this type of content disrupting public compliance with the new regulations and to prevent the rise of incivility (Su et al. 2003)(Kim 2018).

Due to the scale of social media platforms, processing the sheer quantity of posts is a difficult challenge which would require an economically unviable workforce for manual moderation. This is where automated techniques such as automatic text analysis can provide great utility. Twitter reported in April of 2019 on their general content moderation: “38% of abusive content that’s enforced is surfaced proactively to our teams for review instead of relying on reports from people on Twitter”. For COVID-19, this means that less people needed to see the potentially harmful tweets before action was taken. This would means these tweets have less potential influence on the population which is a huge improvement from relying on user-reports.

Matriculation number: 200007413

One major focus of COVID-19 misinformation which proliforated during 2020 was conspiracy theories relating to the 5G network activating the virus (Shahsavari et al. 2020). Bussiness insider reported that in the UK alone, 77 cellphone towers were destroyed due to these conspiracy theories (Hamilton 2020). Shahsavari's study attempted to develop a pipeline for automatically detecting conspicacies on social media and in the news. They used automatic text analysis to scrape for topical analysis of reports to gain related posts and then for relationship extraction and named entity recognition. The entire process is is automated and so artices and posts can be analysed much more rapidly than with manual human investigation. Text analysis is an essential part of this process as it.

Figure 2: YouTube's automatic banner below the video player which is detected to relate to COVID-19. Those which are determined as potentially harmful were just removed.

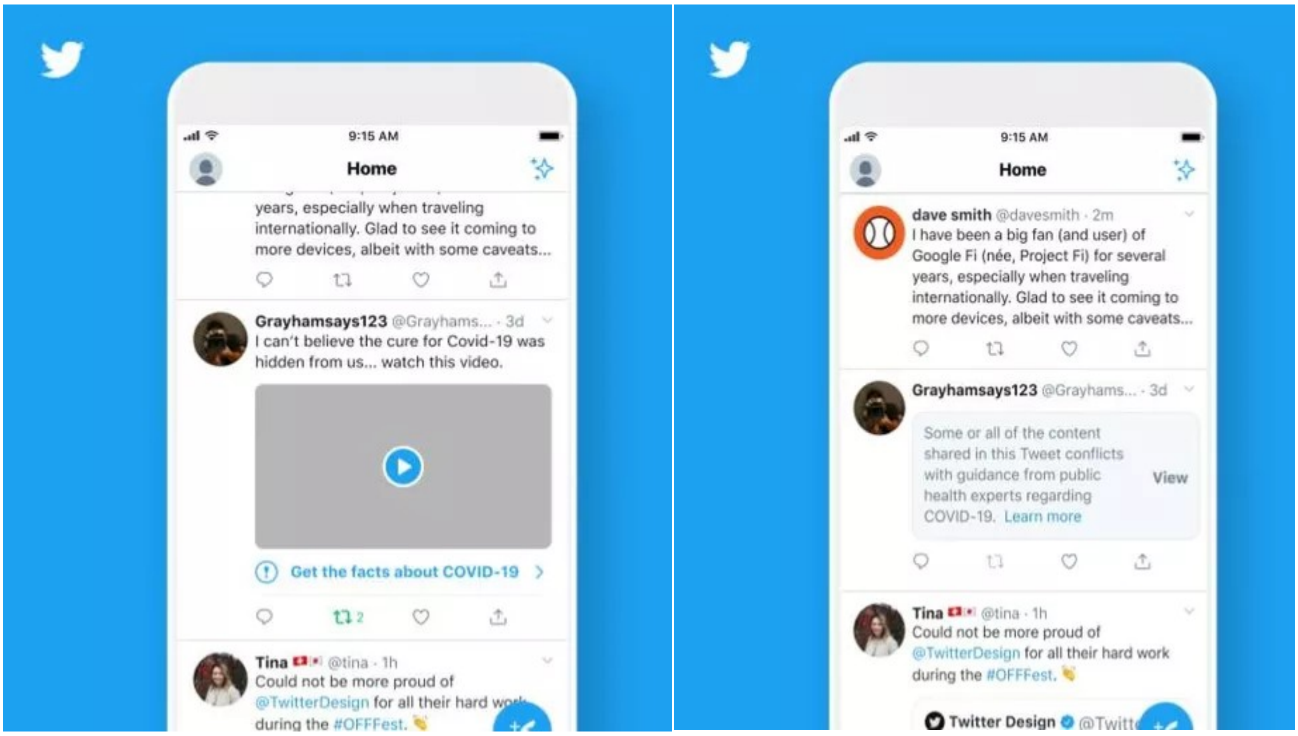

Figure 3: Twitter's automatic flagging system. Disputed claims are labeled (as on the left) when their propensity for harm is moderate and a warning added (as on the right) when propensity for harm is severe. Potentially harmful misleading information with moderate risk is labeled too or removed if a severe risk is determined.

Social media companies have used feature analysis to detect when a piece of content is relating to COVID-19. The use of this is to flag these posts for human moderation as it is more difficult to detect misinformation automatically which shows another downfall of feature analysis. As language has many nuances, it is very difficult for an automatic tool to cross reference what a piece of text is saying with facts – especially when the facts are not obvious in the first place. Social media corporations, however, have used these tools to tag these posts which warn users of the content that the post is potentially misleading and links them to official guidance from trusted sources.

However, these systems are not perfect. Automatic text analysis requires a large sample of misinformation to determine if new posts should be flagged. Brennen et al. (2020) analyzed a sample of 225 pieces of misinformation and compared the action taken by each social media platform. "On Twitter, 59% of posts rated as false in our sample by fact-checkers remain up. On YouTube, 27% remain up, and on Facebook, 24% [..] remains up without warning labels".

Virtual Patient Risk Assessments

As has been made clear over the crisis, COVID-19 is a serious threat to society. As we have learned more about the virus, there has been a focus from governments on social-distancing and self-isolation when you present symptoms. When the virus first was labeled as a serious threat, the key symptoms and risks was relatively unknown and so any way to gain new knowledge was practically useful. Obeid et al. (2020) attempted, therefore, to develop ways of determining how likely it is that a patient has the virus based on self-reported symptoms from virtual visits from their doctor. They used automatic text analysis to build a feature set of patients self-reported symptoms which was then put through a convolutional neural network which could output the likelihood of them having a positive test result. Not only this, they could back-propagate the test results the patients then actually received to improve the algorithm and to gain new information on the virus. This effort had a real-world impact fighting the virus. Their results found certain types of symptoms, such as a loss of smell and taste were surprisingly common in those who tested positive. This lead to the addition of these symptoms to the official list of key symptoms to look out for.

Moreover, the study's output helped the vast number of patients needing tests to be prioritized. At the start of the virus, the capacity for testing was extremely low. It took a long time before wide-spread testing was available and so with a tool like this, those who were flagged as most likely to test positive could be tested first and, even if those who were deemed a high risk for the virus could not get a test, more focus could be applied on making sure they know they need to self-isolate.

Matriculation number: 200007413

Tracking Public Perception

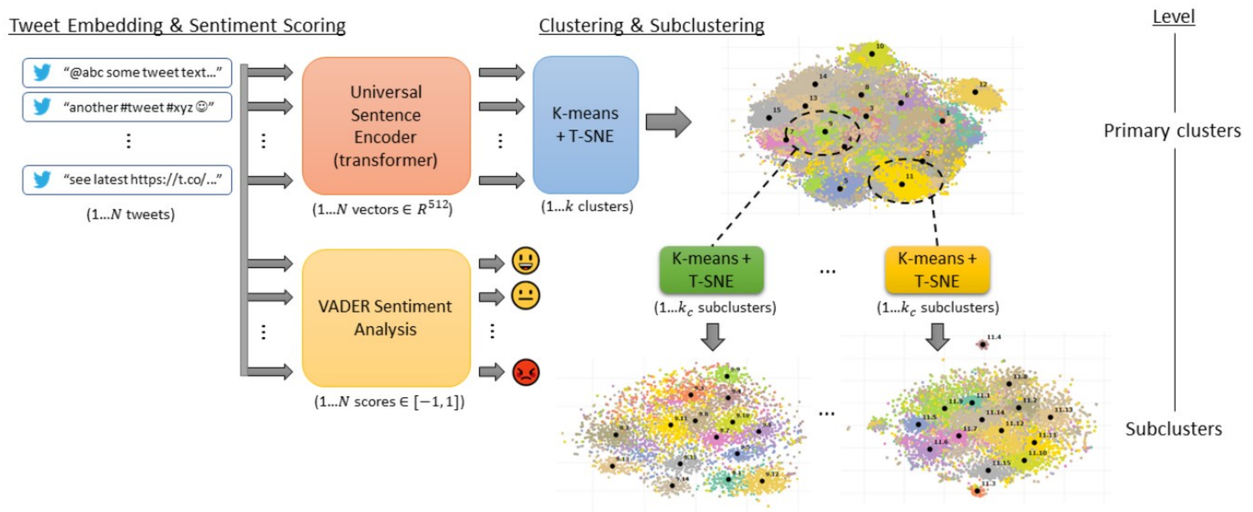

Figure 4: Sanders et al. (2020) - Showing how VADER sentiment analysis was used for the live-tracking of public perception of mask-wearing preventative measures for COVID-19

With new guidelines, comes inevitable reaction from the populace on social media. New studies have been conducted during the COVID-19 pandemic to track the changes in perceptions to governments and their preventative measures. One such study focused on Twitter: Sanders et al. (2020: pre-print) took a database of 1 million tweets from during the pandemic. Figure 4 shows how VADER sentiment analysis was employed to organize tweets relating to mask wearing into their level of positive/negative sentiment to the preventative measure of mask wearing. This allowed the study to follow the reaction of the mask-wearing guidance in real time through the automated pipeline. As this was a new study, it's results and methodology cannot be relied on, however, this shows how sentimental text analysis has practical uses in regards to COVID-19. Other studies have also been conducted to track perceptions of COVID-19 related issues using automatic text analytics tools which can pick out particularly common themes in a dataset. These themes could then be fed into sentiment analysis to gauge the common positive, negative and neutral themes discussed in relation to the pandemic (Aslam et al. 2020)

Potential Guidance for Political Leaders

Continuing on from the previous section, the study by Sanders et al. (along with some others), have shown the practical utility of automatic text analysis. This is because with automated pipelines, real-time information can be gathered from social media. Due to the constantly changing nature of a novel COVID-19 pandemic; unprecedented measures have been taken around the world by governments with lock-downs and mandatory mask-wearing in shops and indoors public spaces. The tools created in these sorts of studies could have been used by the government to keep up with the overall public perception of their guidance and actions. This would allow for better focus on particular concerns of the public to prevent push-back on further measures when they change and to see what measures have the biggest impact on the populace.

Real-time analysis of mass social media posts does, however, bring up some particular issues with text analysis. As these studies have used automatic text analysis for categorization of tweets, it is difficult to determine the biases of the posts the algorithm determines as relating to COVID-19. It could feasibly be true that only certain references are picked up by the algorithm and so other concerns are ignored. If governments are using this as a guide for what to focus on then this could potentially lead to disregard to important issues that just get ignored by the algorithm if the same used to train the model doesn't lead to new types of issues to be detected. As things keep changing, new issues are brought up often and because automatic text analysis struggles to detect these new issues as relevant.

Matriculation number: 200007413

Bibliography

Griffith J. et al. (2003)
SARS and the Modern Day Pony Express (the World Wide Web)
Academic Journal of Roentgenology 2003; 180: 1736
https://www.ajronline.org/doi/10.2214/ajr.180.6.1801736

Wang Y. et al. (2019)
Systematic Literature Review on the Spread of Health-related Misinformation on Social Media.
Social Science & Medicine 2019; 240: 112552
https://www.sciencedirect.com/science/article/pii/S0277953619305465

Su F. et al (2018)
Uncivil and Personal? Comparing patterns of incivility in comments on the Facebook pages of news outlets.
New Media & Society 2018; 20: 3678-3699
https://journals.sagepub.com/doi/full/10.1177/1461444818757205

Kim J. (2018)
Online incivility in comment boards; partisanship matters – but what I think matters more.
Computers in Human Behaviour 2018; 85: 405-412
https://www.sciencedirect.com/science/article/pii/S0747563218301791

Gasca D (Twitter Senior Director) (2019)
A healthier Twitter: Progress and more to do
Official Twitter Blog Post; 2019
https://blog.twitter.com/en_us/topics/company/2019/health-update.html

Shahsavari, S. et al. (2020)
Conspiracy in the time of corona: automatic detection of emerging COVID-19 conspiracy theories in social media and the news
J Comput Soc Sc 2020
https://doi.org/10.1007/s42001-020-00086-5

Hamilton I. (2020)
77 cell phone towers have been set on fire so far due to a weird coronavirus 5G conspiracy theory
Business Insider
https://www.businessinsider.com/77-phone-masts-fire-coronavirus-5g-conspiracy-theory-2020-5

Sanders A. et al (2020: pre-print)
Unmasking the coversation on masks: Natural language processing for topical sentiment analysis of COVID-19 Twitter discourse
medRxiv 2020: pre-print
https://doi.org/10.1101/2020.08.28.20183863

Matriculation number: 200007413

Aslam F. et al. (2020)
Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak.
Hamanit Soc Sci Commun 2020: 7, 23
https://doi.org/10.1057/s41599-020-0523-3

Obeid J. et al. (2020)
An artificial intelligence approach to COVID-19 infection risk assessment in virual visits: A case report
Journal of American Medical Informatics Association 2020; 27, 8: 1321-1325
https://doi.org/10.1093/jamia/ocaa105

Brennan J. et al. (2020)
Types, sources, and claims of COVID-19 misinformation
Reuters Institute 2020; 7: 3-1
http://www.primaonline.it/wp-content/uploads/2020/04/COVID-19_reuters.pdf