Detecting Covid-19 misinformation on social media

Jason Michaud, *Bryant University, jmichaud1@bryant.edu*
Suhong Li, *Bryant University, sli@bryant.eud*

Abstract

There have been many studies conducted over the last few years that have attempted to uncover the impacts of the COVID-19 pandemic. One of the largest areas of concern with COVID-19 is misinformation, as it is a novel virus that many report on, even if unqualified to do so. This study aims to predict whether a Tweet can be classified as misinformation, and then analyze the differences between Tweets that are labeled as either misinformation or not by this model in the three countries (The United States, UK, and India). Machine learning models were created and validated using the CovidMis20 dataset as a training set. The result show that for the United States, there is strong evidence that the model worked at predicting fake news. Hydroxychloroquine showed up in both top hashtags and topic modeling for fake news. In the UK, the real news dataset tended to be more general and objective about COVID-19 than predicted fake news. For India, there was not much evidence that the model was effective for this country.

Keywords: misinformation, COVID-19, machine learning, topic modeling

Introduction

The COVID-19 pandemic has turned life upside down across the globe, and impacts can be felt to this day more than three years later. With constantly changing guidelines, various accounts of information coexisted, some of these being credible and others not. COVID-19 misinformation has been an area of concern for many, especially with social media, as viral messages can spread extremely quickly. The purpose of this study is to first develop a machine learning model to predict whether a Tweet is misinformation or not in three countries (US, UK and India), followed by identifying/comparing dominant topics of misinformation emerging in those three countries.

Literature Review

With the COVID-19 pandemic having been in existence for more than two years now, there have been many studies conducted and published already on the topic. Yu et al. (2022) examined the impact of COVID-19 tweets with negative sentiment and found that negative sentiment leads to a higher number of favorites and replies. Rustam et al. (2021) developed various machine learning models to predict the sentiment of COVID-19 tweets and reached a high accuracy of 93%. Another study by Cuomo et al. (2021) looked at a geospatial analysis of COVID-19 tweets and used a combination of machine learning models and geospatial analysis techniques to search for a relationship to future case counts. The study found that there is some relationship between the two. In addition, Hswen et al. (2021) investigated anti-Asian sentiment with covid-19 Tweets in March 2020 and found that there was a rise in anti-Asian sentiment.

Bonnevie et al. (2021) found that vaccine opposition increased on Twitter once COVID-19 started to become a larger issue.

A few studies focus on identifying dominant topics in covid-19 tweets using Latent Dirichlet Allocation (LDA) technique. For example, Jafarzadeh et al. (2021) collected tweets during a major Covid-19 event in New Zealand and identified main topics discussed including health matters, government response and measures, and political and economic impact. Another study (Xu et al., 2020) used 1.9 million covid-19 tweets and identified such themes as "Updates about the number of COVID-19 cases", "COVID-19 related death", "Cases outside China", "Outbreak in South Korea", "Early signs of the outbreak in New York City", "Diamond princess cruise", "economic impact", "Preventive measures", "Authorities", and "Supply Chain". A third study that utilized LDA Topic Modeling was done on vaping during the COVID-19 pandemic (Chen et. al, 2021). The main topics found in this study include "Linkage between vaping and risk of COVID-19 infection", "Vaping pneumonia is the origin of COVID-19", "Vaping and spread of COVID-19", "Vaping regulation", "Calling for quitting vaping", "Protecting youth", "Vaping-related lung injury", and "Sales information."

There are other studies focusing on the impact of social media data on covid-19 cases. For example, Jafari et al. (2020) used tweet frequency and Google Search on COVID-19 symptoms to predict the occurrence of COVID-19. Ayyoubzadeh et al. (2020) used Google Trends for various COVID-19-related terms to predict Covid-19 cases.

Studies on COVID-19 Misinformation

Many studies have attempted to uncover the effects of misinformation throughout the COVID-19 pandemic. World Health Organization stated that "In the first 3 months of 2020, nearly 6000 people around the globe were hospitalized because of coronavirus misinformation". This is significant, and it really highlights why misinformation is such a pressing issue with a novel virus like COVID-19. To combat this, the WHO and the United Kingdom started the "Stop the Spread" campaign, and this campaign was really an attempt to have a centralized distribution of COVID-19 information in the United Kingdom in order to avoid any unnecessary consequences.

Snyder et al. (2020) examined misinformation in COVID-19 crowdfunding campaigns and found 208 campaigns were identified for unproven treatments for COVID-19. The most common among these treatments were immune system boosters and dietary supplements. Actual medicinal treatments were included in 15 of these campaigns, with one of these treatments being Hydroxychloroquine. These crowdfunding campaigns reach thousands of people, and this really shows how quickly misinformation can spread. Preston et al. (2021) conducted a study in 2020 to uncover the role of emotional intelligence in detecting fake news on Facebook. 87 participants were included to see whether they could detect fake news. One of the main intriguing conclusions from this study was that the more emotionally intelligent a person was, the less likely they would fall for fake news or misinformation.

On top of just studying misinformation with COVID-19, studies have been done to predict misinformation using machine learning methods. The first of these was a study in 2021 that aimed to predict fake news in COVID-19 Tweets. While the authors do not mention how many tweets they used, they did achieve a weighted average f1-score of 98% (Bangyal, et al., 2021). Rather than using tweets, another study (Hansrajh et al., 2021) used news articles to predict real and fake news. They used blending ensemble methods based on both the Liar and ISOT datasets. The blending ensemble methods across different models ended up being more successful for the ISOT dataset with an f1-score of .984. For the Liar dataset, f1-score is only .682.

<center>**Research Methodology**</center>

Data Collection Methodology

Tweets were collected between March 2020 and May 2022 using keyword covid-19 and over 1.6 billion tweets were collected. The study will focus on tweets in the three counties (US, UK and India). User location was used to identity the original of tweets. The user location is self-reported field, not all Twitter accounts have a user location. In addition, users may enter a fake location. Therefore, the study will not represent all tweets collected during this study period and only focus on the tweets that is associated with a valid user location in the three countries. For each country, a list of state and province names was needed to capture the most data. For the United States, all tweets containing either the state name, province name, state abbreviation, or province abbreviation were captured. For the United Kingdom, all countries in it were captured, as well as the user location of "UK". For India, all 28 states were captured. Finally, for all three countries, each respective country name was a filter word as well, and the data was saved in Parquet format for further analysis. The total tweets for the Unites States, UK and India are 218,249,384, 43,407,501, and 28,910,111 respectively.

Machine Learning Methodology

Once in Parquet format, the tweets were then uploaded to a separate, more powerful cluster run by the Massachusetts Green High Performance Computing Center (MGHPCC) for machine learning and analysis purposes. Access to this cluster was provided by the CAREERS Cyberteam Project and the University of Rhode Island.

The next step was to identify a training set that a model could be developed from to predict whether a tweet was misinformation or real news. After much exploration, the CovidMis20 dataset was chosen for this. This dataset "includes around 1,375,592 tweets from February to July 2020" (Mulahuwaish et al., 2022). To train their model to label the Tweets, the authors used a predetermined list of websites obtained from MBFC (Media Bias Fact Check) and then attached these labels to the tweets. They then created a machine learning model of their own and labeled each Tweet as fake or real.

Once the CovidMis20 dataset was uploaded to the Unity Cluster, it was then cleaned before training any model. The text cleaning pipeline is as follows. Text was assembled into documents, documents into sentences, sentences into tokens, tokens into lemmas, lemmas into normalized text free of any special characters, and finally the data was passed through a finisher. The result of this is a clean piece of text for every Tweet in the CovidMis20 dataset.

Once clean, both a Logistic Regression and Naïve Bayes model were utilized to develop the best performing model to apply to this project's tweets. The output for both the Naïve Bayes model and Logistic Regression can be found below in Figures 1 and 2. Figure 1 is the model training results for Naïve Bayes, and Figure 2 is the model training results for Logistic Regression.

As it can be seen in Figures 1 and 2, models trained from the CovidMis20 dataset performed very well. Although the lift and gain curves look very similar, the Logistic Regression is clearly the better model when the f1-scores are compared. For real news, the f1-score is 0.02 higher for the Logistic Regression than Naïve Bayes. For misinformation, the f1-score is 0.03 higher for the Logistic Regression than for Naïve Bayes. Therefore, the Logistic Regression was saved and chosen as the final model to apply to the unlabeled Tweets for each country.

Data Analysis

For the purposes of this project, data was analyzed using Spark on the MGHPCC Cluster. The four types of analysis performed were top hashtags, word clouds, top users tagged, and LDA Topic Modeling. These four types of analysis were run in two separate iterations for each country, one for predicted real news, and one for predicted fake news. These four types of analysis mainly explore the text content of each Tweet to determine key differences between predicted real news and predicted misinformation.

Word Cloud Analysis

United States Word Clouds: In Figure 3, the word cloud for the United States fake news dataset is displayed. After stop words and COVID-19-related words are removed, there are some key topics that can be seen. Some of the most frequent words here include China, tested, media, state, spread, last, and virus. It makes sense that China showed up here, as there is much conspiracy and misinformation on how they handled the virus, as this is the origin country of COVID-19. It is also interesting that the terms tested and state showed up, as this implies a statistics-heavy area in this dataset as well that was predicted as misinformation.
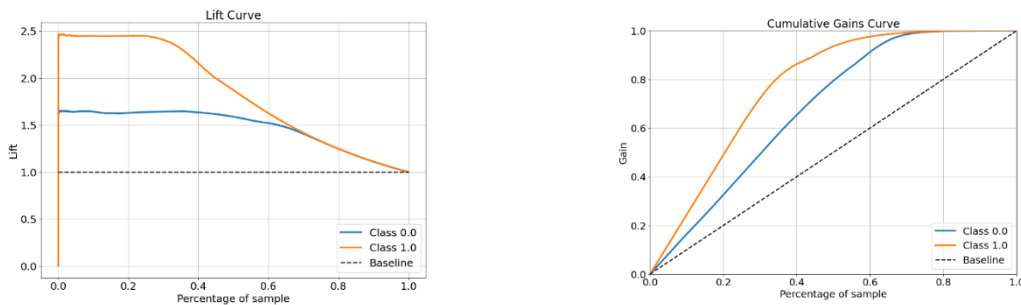
Label	Precision	Recall	F1-score	Support
0 (real)	0.90	0.93	0.91	245811
1 (fake)	0.89	0.85	0.87	165930

Figure 1: Naïve Bayes Model Training Results
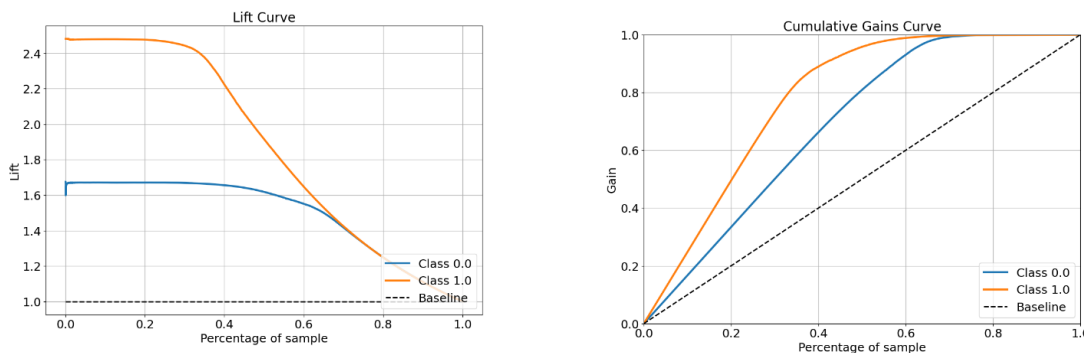
Label	Precision	Recall	F1-score	Support
0 (real)	0.91	0.96	0.93	245811
1 (fake)	0.93	0.86	0.90	165930

Figure 1: Logistic Regression Training Results

The word cloud for United States real news can be seen in Figure 4. Although there are many similar terms here to the fake news set, there are also some key differences. The main difference is that the terms CDC, white, and house show up here. This implies a news and politics-heavy dataset, as white and house showing up together may imply the White House. This would make sense too, as there was a presidential election that happened during data collection. It is also intriguing that CDC showed up here, as they set official COVID-19 guidelines in the United States. This implies that the model may actually be working for this country's dataset, but more analysis is needed before that can be said for sure.

Figure 3: United States Word Cloud Fake News

Figure 4: United States Word Cloud Real News

United Kingdom Word Clouds: For the United Kingdom, there were much more clear-cut results with the word clouds that can be related back to the topic of misinformation. Starting off, the fake news word cloud for the United Kingdom is displayed in Figure 5. The top words here include Boris, Johnson, deaths, government, lockdown, care, time, staff, and died. As can be seen by the word size, one of the top areas of interest here in the United Kingdom dataset predicted as fake news is Boris Johnson.

It does make sense that Boris Johnson, the ex-prime minister of the United Kingdom does show up here, as there was a scandal involving him and other officials breaking their own COVID-19 guidelines by attending a party. This event created huge traffic on social media and is one explanation for the model behaving this way, as with such viral events there are always conflicting reports that exist. Another explanation is simply that the CovidMis20 dataset only performs well for United States based news topics.

Figure 5: United Kingdom Word Cloud Fake News

The United Kingdom real news word cloud can be seen in Figure 6. The top words include deaths, many, know, need, last, year, day, children, omicron, and work. Here, with deaths as the top word, it seems like much more diverse topics exist rather than the political topics in the fake news word cloud. More of a statistical picture and a focus on guidelines seems to be what is going on in this figure. It is interesting that there is such a difference between the fake news and real news word clouds for the United Kingdom. The fake news one has more of a political focus while the real news one has much more objective words, and meaningful insights can be derived here.

Figure 6: United Kingdom Word Cloud Real News

India Word Clouds: For India, the word clouds seemed to imply that the model may have not been the best for data of this nation. The fake news word cloud for India is displayed in Figure 7. The most frequent words are larger, and for this word cloud, Delhi, vaccine, govt, government, and minister are among the top words. While this does imply a discussion of vaccines predicted as fake news, it also implies that much country-specific news about India was categorized as fake news. While there may be a legitimate claim to fake news here, it is most likely the case that the CovidMis20 dataset and MBFC that it was trained on did

not have much exposure to India-specific COVID-19 news.

Figure 7: India Word Cloud Fake News

Figure 8: India Word Cloud Real News

For predicted real news, India did not have much different results than the fake news word cloud. The results are displayed in Figure 8. While vaccine was a key topic here, there were some differences to the fake news word cloud. The top words here included oxygen, students, help, today, please, and need. This seems to imply much more of a conversational atmosphere as compared to the news heavy atmosphere of the fake news word cloud. Between the two-word clouds of India, there was definitely some divide in the prediction of the model trained from the CovidMis20 dataset, but its accuracy for this country cannot be determined from this type of analysis.

Top Hashtags Analysis

For each country, the top hashtags were found by simply just counting them by frequency in each Tweet. Some generic COVID-19 hashtags were removed from each dataset prior to this analysis, as these mainly had to do with how the original dataset was collected and did not offer much insight here.

United States Top Hashtags: For the United States, the top 10 hashtags for predicted fake news are displayed below In Table 1. When Trump was President and active on Twitter, many of his Tweets were rooted in opinion and not fact, so it is no surprise that #Trump is the second most frequent hashtag. #FoxNews and #China showing up is also interesting, as Fox News is known to be more on the conservative side, and something related to China may be being flagged as fake news.

The most interesting hashtag that showed up here is #Hydroxycloroquine. Hydroxychloroquine is a drug used to treat Malaria and was widely pushed by President Trump to treat COVID. This was proven ineffective according to Medlineplus.gov and it makes sense that this is showing up in the fake news dataset (Hydroxychloroquine, 2023).

For predicted real news, there is a slightly different picture being painted, and the top hashtags can be found in Table 1 also. While #Trump is still among the most frequent terms, #FoxNews is knocked down a bit on the list. #SmartNews takes over the number one spot, and this is a hashtag used on Twitter referencing the news app Smart News. There are also some COVID-related terms that showed up here such as #WearAMask and #vaccine, and this makes sense in terms of being flagged as real news. #DeathSantis refers to Governor DeSantis of Florida, and it is people who oppose his decisions who take to Twitter using this hashtag. This makes sense that it shows up, as many are just expressing their opinion against him and it would not be flagged as fake news or misinformation.

Table 1: United States Top Hashtags Fake vs Real News

Fake News		Real News	
Hashtag	**Count**	**Hashtag**	**Count**
#China	26281	#SmartNews	145539
#Trump	23460	#Trump	108622
#FoxNews	22242	#1	96336
#LongCovid	21379	#Florida	71677
#lockdown	16332	#FoxNews	69620
#mask	13994	#vaccine	57911
#Hydroxychloroquine	13906	#WearAMask	56830
#Iran	13747	#DeathSantis	52290
#SmartNews	13659	#China	51749
#Texas	12414	#OANN	44911

United Kingdom Top Hashtags: For the UK, the top 10 hashtags for fake news can be found in Table 2. Some of these are quite interesting and require a bit further exploration to uncover meaning. The first one of these is #BorisJohnson, which is the third most frequent hashtag in the dataset. Boris Johnson is the ex-prime minister of the United Kingdom, and there was a party scandal that he was a part of mentioned earlier in this paper. It makes sense that he shows up here, as he is a polarizing figure for many.

It is also interesting that #LongCovid shows up here. While this is not necessarily fake news, long COVID is a term that was not coined until well into the pandemic, so the CovidMis20 dataset may not be accustomed to this term. #Brexit also shows up, but with MBFC being United States based, this may just be because the training dataset has not seen this term. #PMQs stands for Prime Minister Questions, and it is interesting that this shows up as well relating back to Boris Johnson. Lastly, #Marr seems to be out of place, as no real meaning could be found for this hashtag.

The top 10 hashtags for predicted real news for the United Kingdom can also be found in Table 2, and it is a very different picture than the fake news dataset. These are all COVID-19 related, and the difference between the fake and real predicted groups is quite interesting here. In the fake news set, there was a lot of political talk, but there is none in the predicted real news hashtags. This is evidence that the model may have worked to a certain extent here, as there is a clear-cut distinction between predicted groups.

Table 2: United Kingdom Top Hashtags Fake vs Real News

Fake		Real	
Hashtag	**Count**	**Hashtag**	**Count**
#NHS	50399	#giveaway	13694
#lockdown	33055	#LongCovid	11218
#BorisJohnson	25667	#Health	10779
#LongCovid	23663	#SARSCoV2	9669
#UK	23141	#vaccine	9561
#Brexit	18094	#pandemic	7951
#vaccine	13528	#SocialDistancing	6247
#PMQs	13505	#lockdown	5692
#ExcludedUK	10853	#China	5275
#Marr	9113	#mentalhealth	5211

India Top Hashtags: In Table 3, the top 10 hashtags in the India fake news dataset by frequency are displayed. While most of these relate to either news or are India-specific regions, #Blood and #Neet had to be explored further with a search on Twitter. It seems that #Blood is related to COVID, and it is interesting that this is the top hashtag in documents predicted as fake news for this dataset. When searched, it was found that #Neet is the National Eligibility and Entrance Test. This is something that medical students have to take in India, but it does seem that there are some unrelated uses of this hashtag on Twitter too. With such location-specific results, it is tough to say whether the model works for this dataset, or if it is just classifying words and terms, it has never seen as fake news automatically.

Going into the real news hashtags for India, it is a very similar picture to the fake news ones. The top 10 hashtags for India real news can also be found in Table 3. #Blood is also the top term here. Again, there are also many India-specific hashtags such as #Odisha, #Rajasthan, and #Hyderabad. Similar to the India word clouds, there is no clear picture that seems to be being painted by this dataset. It seems that the model may not be working the best, but this makes sense because Indian news may not have been a major part of the CovidMis20 dataset.

Top Users Tagged

Similar to top hashtags, top users were collected by counting the frequency of a user being tagged in the text of a tweet by specifically looking for the '@' symbol. Stopwords did not necessarily play an impact here, as these are just usernames in the table and not a body of text needing to be filtered.

United States Top Users Tagged: For the United States, the top users tagged in the fake news dataset can be found below in Table 4. There are two key users to be highlighted here, but unfortunately, there was also a lot of noise in the dataset. The two users to be covered here are @realDonaldTrump and @drdavidsamadi. @realDonaldTrump is former President of The United States Donald Trump, and it was expected that he would show up here. His Tweets created much traffic before he was banned from Twitter, and these Tweets were not always rooted in fact.

This is a good sign for the model that he showed up here. @drdavidsamadi is a doctor named David Samadi, and also a COVID critic which can be found by examining his Tweets. He was banned for a week once for spreading the idea that hospitals were faking numbers at the time, as well as other similar criticisms of COVID. Him being a top user in the fake set makes sense with the goal of predicting misinformation.

Table 3: India Top Hashtags Fake vs Real News

Fake		Real	
Hashtag	**Count**	**Hashtag**	**Count**
#Blood	38233	#Blood	76935
#Maharashtra	25112	#IndiaFightsCorona	22843
#lockdown	24977	#Odisha	19587
#SOS	20944	#COVID19India	17765
#BREAKING	17996	#BREAKING	16180
#Neet	15796	#Rajasthan	15649
#IndiaFightsCorona	14937	#BiharFightsCorona	14770
#China	14894	#Hyderabad	14475
#WATCH	14060	#China	14331
#Kerala	13967	#SOSIYC	13990

For the real news dataset for the United States, the top users tagged can also be found in Table 4. Here, there are a few interesting users as well. It is intriguing that @realDonaldTrump is also the top user in this dataset, but this makes sense. As mentioned before, his Tweets are extremely viral, and not everything he says is objective misinformation.

It is also interesting that other politicians show up here such as @JoeBiden and @RealCandaceO who is Candace Owens. One final insight here is that many news organizations show up, which is expected if they are doing their jobs and objectively reporting factual information.

Table 4: United States Top Users Tagged Fake vs Real News

Fake		Real	
Username	**Count**	**Username**	**Count**
@realDonaldTrump	231903	@realDonaldTrump	1985465
@jsolomonReports	225312	@kylegriffin1	1320935
@drdavidsamadi	190694	@JoeBiden	697775
@CNN	182245	@DrEricDing	640261
@Reuters	161711	@RealCandaceO	524321
@DrEricDing	143291	@CNN	511835
@BNODesk	118421	@nytimes	511416
@OccupyDemocrats	108858	@catturd2	487951
@JoeBiden	108405	@ProjectLincoln	471551
@NBCNews	104084	@funder	466730

United Kingdom Top Users Tagged: For the United Kingdom, there was a lot of overlap between tagged users in the real and fake datasets, so both will be analyzed in one paragraph. The top users tagged for both fake and real news can be found in Table 5. There are two influential users who showed up in the fake dataset that are interesting, and no real standout users for the real news dataset. The two interesting users in the fake set are @BorisJohnson and @PeterStefanovi2. Boris Johnson is the ex-prime minister of the United Kingdom, and Peter Stefanovic is a lawyer, vlogger, and filmmaker as per his Twitter bio. There is much controversy surrounding Johnson, and it is no surprise that he was tagged many times in the fake news dataset. Stefanovic came out of nowhere though, and his high presence is actually due to a Tweet about Boris Johnson. He is calling out Johnson for lying and pushing false information in a video, so it

makes sense that Stefanovic is a top user tagged for this dataset as well. This video actually achieved over 45 million views to date and is the pinned Tweet on Stefanovic's account.

Table 5: United Kingdom Top Users Tagged Fake vs Real News

Fake		Real	
Username	**Count**	**Username**	**Count**
@BorisJohnson	289886	@doctor_oxford	70019
@SkyNews	195952	@EssexPR	67010
@piersmorgan	180198	@DrEricDing	44633
@doctor_oxford	148946	@realDonaldTrump	40305
@JamesMelville	138613	@njoyflyfishing	38046
@MattHancock	125971	@JamesMelville	35994
@PeterStefanovi2	117503	@sajidjavid	32796
@NHSMillion	112977	@Kit_Yates_Maths	32668
@davidschneider	107593	@MichaelPSenger	32621
@chrischirp	98323	@SkyNews	30045

India Top Users Tagged: For India, no key purveyor of misinformation could be identified upon investigating the top users in the dataset. Top users for both fake and real news can be found in Table 6 below. There is much overlap between the datasets here, as many political figures and news agencies are showing up in both. An example of this is the Prime Minister of India, Narendra Modi, showing up as the top user in both datasets. This is most likely due to the model not understanding India-specific terms. Although all Tweets are in English, it is likely that India-specific regions and states threw off the model, which is why there are no key results in this section.

Table 6: India Top Users Tagged Fake vs Real News

Fake		Real	
Username	**Count**	**Username**	**Count**
@narendramodi	352797	@narendramodi	271318
@ANI	234427	@ANI	176540
@PTI_News	147813	@AskAnshul	136024
@AskAnshul	126729	@RahulGandhi	111004
@OpIndia_com	114560	@ndtv	106389
@RahulGandhi	93961	@OPIndia_com	100768
@NorbertElekes	75104	@PMOIndia	98364
@spectatorindex	56184	@PTI_News	94442
@INCIndia	52999	@NorbertElekes	92346
@ArvindKejriwal	48042	@FaheemYounus	80878

Topic Modeling

Topic Modeling was one of the main types of analysis performed for this project and was where the most meaningful results came from. Specifically, LDA Topic Modeling was performed in PySpark in the HPC Environment using the PyLDAvis library. Topic modeling was performed for each country for all tweets predicted as either fake or real respectively. In total six topic modeling visualizations were run in this study. The output for fake tweets in the United States is shown in Figure 9 as an example and the output for the other topic modelling analysis are ignored due to size limitation. From Figure 9, it can be seen that LDA topic modeling produced a visual with numerical topics. For ease of reading, topics were given a name based on the terms in them and combined if similar, and results can be found in Table 7.
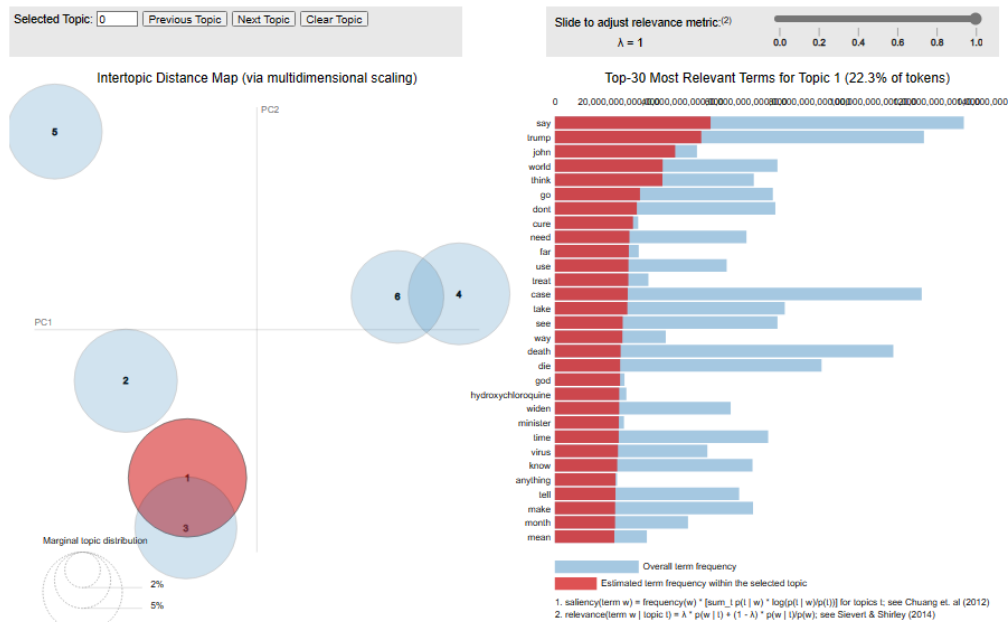
Figure 9 LDA Topic Modeling Example United States Misinformation

Table 7 shows that for the United States, when examining the different fake news topics for this country, some key terms include take, treat, Hydroxychloroquine, refuse, vaccinate, and lab. Hydroxychloroquine is a drug used to treat Malaria and was widely pushed by President Trump to treat COVID. This was proven ineffective according to Medlineplus.gov and it makes sense that this is showing up in the fake news dataset (Hydroxychloroquine, 2023). Lab refers to the conspiracy theory that COVID-19 was engineered in a lab, and this also makes sense to be predicted as misinformation.

For the United Kingdom, there is much political talk in the topic modeling with Boris being among the top terms for misinformation, referring to Boris Johnson. While Boris Johnson is a controversial figure, this may also be explained by the model not having much United Kingdom specific data. Finally, for India, the topic modeling raised concerns about the accuracy of the machine learning model for this country. Much location-specific data such as 'Delhi' and 'Maharashtra' was predicted as misinformation, and this points to the model being ineffective and unfamiliar with these terms.

Conclusions

For the United States, there is strong evidence that the model worked at predicting fake news when all four analysis types are examined together. Hydroxychloroquine showed up in both top hashtags and topic modeling for fake news, and this is a very promising result. It is also promising that the word lab showed up in the fake news topic modeling as well, as this relates to the conspiracy theory of COVID being engineered in a lab. Similar to the other countries, however, there was a lot of overlap between fake and real news for the United States, so there is no way to truly tell how well the model has performed.

Table 7: Topic Modeling Results

Country	Label (Fake or Real)	Topic Name	Terms
United States	Fake	1) Covid Symptoms/ Work from Home	Long, symptom, ill, work, worker, home, relief, night
		2) Politics	Test, Trump, Biden, Fauci, Kamala, lab
		3) Covid Treatment	Say, Trump, world, cure, Hydroxychloroquine, treat, case, take, vaccinate, die, death, risk, refuse, make, believe
		4) Cases/Lockdown	Social, lockdown, mask, free, distance, death, government, south, level, immunity
United States	Real	1) Mask mandate and social distancing	Test, mask, school, mandate, child, spread, require, flu, kid, stay, home
		2) Covid Statistics	Case, death, report, state, break, death, million, die
		3) Politics	Trump, president, CDC, Biden, Fauci, data, lie, force, bill, republican, money, vote
United Kingdom	Fake	1) Covid Treatment	Rate, government, high, infection, dose, death, variant, lockdown, system, failure, test, study, vaccination, Trump, say
		2) Politics	Public, government, rule, break, Boris, minister, restriction, hospitalization, death, care, home, omicron
		3) Children	Child, support, worker, social, distance, flu, age, fight, die, jab
United Kingdom	Real	1) Covid Spread	Test, share, house, omicron, variant, die, need, care, take, patient, vaccinate, bad, spread, child, party, minister
		2) General Covid	Case, update, death, report, daily, united, important, social, population
India	Fake	1) Covid Statistics	Case, death, report, indie, rise, record, total, number, day, vaccine, Delhi, Kerana, Maharashtra
		2) Feelings	Need, help, please, care, patient, Delhi, oxygen, government, doctor, warrior, police, speedy, wish, recovery, vaccine, minister
India	Real	1) Statistics	Case, death, report, vaccine, day, number, total, omicron, million, update
		2)High Spirits/Feelings	Fight, thank, contact, say, today, lose, one, prime, minister, get, well, please, help, mask, safe, need, please, blood, plasma
		3) Students	Exam, student, oxygen, bed, postpone, conduct, icu, ventilator

The United Kingdom was the country with the largest difference between predicted fake news and real news. This is due to the topic of Boris Johnson and politics showing up in the fake news dataset across all types of analysis. The real news dataset tended to be more general and objective about COVID-19 here as well. This is promising regarding the performance of the model, but it could be due to another factor also. The training dataset was based on MBFC which is United States based, so it is possible that some Tweets that were United Kingdom themed were classified as misinformation simply due to the model not having access to those headlines before.

For India, the conclusion is very short and was alluded to in the analysis sections. There was not much evidence that the model was effective for this country. There was much location-specific language, such as

---

Delhi and Maharashtra in the Tweets, and this could have very well thrown off the model in accuracy. Again, MBFC is a United States based company, and it is likely that the training dataset does not have India-specific headlines. With cases and deaths differing by country too, this is another factor that may have led to there being not many insights for India between fake and real news.

With results that make it seem like the model performed better for some countries than others, there are many steps that can be taken to potentially improve results going forward. First and foremost, a more diverse training dataset can be used. As alluded to throughout the paper, there were signs that the model was limited in accuracy for India and the United Kingdom in the analysis of results. Better results here can be achieved in two ways. The first way is to combine additional training datasets with CovidMis20. This would guarantee a more diverse pool of text for the model to train with. The other way of improving the model's accuracy here is to obtain Tweets from India and the United Kingdom that have been labeled by experts as fake or real. This would solve the issue of country specific news being falsely labeled as misinformation and may offer more intriguing results for the project. As mentioned previously in this section, the United States dataset seemed to have some promising results in the detection of misinformation as evidenced by all of the analysis results. Another future direction for this study is to do a more drilled down analysis of the United States dataset at the state level. With over 200 million Tweets, there is a lot more information in this subset than what was just uncovered in this study. A drilled down analysis of the United States at the state level may be key to uncovering these hidden insights.

One final area of future expansion for this project is to do a network analysis of users. With proper computational power, this would be done through a tool like Gephi. With a network analysis algorithm run for each country, only then could the most influential users truly be discovered. This type of analysis would offer insight into how the misinformation spreads and who the key figures are also. This would be a very interesting area to go deeper into, but it was mainly time that was lacking for it to be completed.

References

Ayyoubzadeh, S. M., Ayyoubzadeh, S. M., Zahedi, H., Ahmadi, M., & Niakan Kalhori, S. R. (2020). Predicting COVID-19 Incidence Through Analysis of Google Trends Data in Iran: Data Mining and Deep Learning Pilot Study. *JMIR Public Health and Surveillance*, 1-7.

Bangyal, W., Qasim, R., Rehman, N. u., Ahmad, Z., Dar, H., Rukhsar, L., . . . Ahmad, J. (2021). Detection of Fake News Text Classification on COVID-19 Using Deep Learning Approaches. *Computational & Mathematical Methods in Medicine*, 1-12.

Bonnevie, E., Gallegos-Jeffrey, A., Goldbarg, J., Byrd, B., & Smyser, J. (2021). Quantifying the rise of vaccine opposition on Twitter during the COVID-19 pandemic. *Journal of Communication in Healthcare*, 12-15.

Chen Lyu, J., Luli, G. K., & Ling, P. M. (2021). Vaping discussion in the COVID-19 pandemic: An observational study using Twitter data. *Public Library of Science*, 1-10.

Coronavirus in the U.S.: Latest Map and Case Count. (2022, April 30th). Retrieved from The New York Times: https://www.nytimes.com/interactive/2021/us/covid-cases.html

Cuomo, R. E., Purushothaman, V., Li, J., Cai, M., & Mackey, T. K. (2021). A longitudinal and geospatial analysis of COVID-19 tweets during the early outbreak period in the United States. *BioMed Central Public Health*, 1-9.

Hansrajh, A., Adeliyi, T. T., & Wing, J. (2021). Detection of Online Fake News Using Blending Ensemble Learning. *Scientific Programming*, 1-7.

Hswen, Y., Xu, X., Hing, A., Hawkins, J. B., Brownstein, J. S., & Gee, G. C. (2021). Association of "#covid19" Versus "#chinesevirus" With Anti-Asian Sentiments on Twitter: March 9–23, 2020. *American Journal of Public Health*, 956-961.

Hydroxychloroquine. (2023, Mar 15). Retrieved from Medline Plus: https://medlineplus.gov/druginfo/me

Jafari, A., MacDonald, B., DeConde, A. S., & Panuganti, B. A. (2020). Predicting COVID-19 Incidence Using Anosmia and Other COVID-19 Symptomatology: Preliminary Analysis Using Google and Twitter. *American Academy of Otolaryngology-Head and Neck Surgery Foundation*, 491-496.

Jafarzadeh, H., Pauleen, D. J., Abedin, E., Weerasinghe, K., Taskin, N., & Coskun, M. (2021). Making sense of COVID-19 over time in New Zealand: Assessing the public conversation using Twitter. *Public Library of Science*, 1-10.

Mulahuwaish, A., Osti, M., Gyorick, K.,, Maabreh, M., Gupta, A., & Qolomany, B. (2022). CovidMis20: COVID-19 Misinformation Detection System on Twitter Tweets using Deep Learning Models,". *14th International Conference on Intelligent Human Computer Interaction (IHCI-2022), Tashkent, Uzbekistan, Oct. 2022*.

Preston, S., Anderson, A., Robertson, D. J., Shephard, M. P., & Huhe, N. (2021). Detecting fake news on Facebook: The role of emotional intelligence. *Public Library of Science*, 1-10.

Rustam, F., Khalid, M., Aslam, W., Rupapara, V., Mehmood, A., & Sang Choi, G. (2021). A performance comparison of supervised machine learning models for Covid-19 tweets sentiment analysis. *Public Library of Science*, 1-6.

Snyder, J., Zenone, M., & Caulfield, T. (2020). Crowdfunding Campaigns and COVID-19 Misinformation. *American Journal of Public Health*, 739-741.

World Health Organization. (2021, April 27). *Fighting misinformation in the time of COVID-19, one click at a time*. Retrieved from World Health Organization: https://www.who.int/news-room/feature-stories/detail/fighting-misinformation-in-the-time-of-covid-19-one-click-at-a-time

Xue, J., Chen, J., Chen, C., Zheng, C., Li, S., & Zhu, T. (2020). Public discourse and sentiment during the COVID 19 pandemic: Using Latent Dirichlet Allocation for topic modeling on Twitter. *Public Library of Science*, 1-7.

Yu, H., Yang, C.-C., Yu, P., & Liu, K. (2022). Emotion diffusion effect: Negative sentiment COVID-19 tweets of public organizations attract more responses from followers. *Public Library of Science*, 1-16.

Zhang, X., Saleh, H., Younis, E. M., Sahal, R., & Ali, A. A. (2020). Predicting Coronavirus Pandemic in Real-Time Using Machine Learning and Big Data Streaming System. *Hindawi Complexity*, 1-7.