Modeling and Visualization of Long-Term Public Opinion on COVID-19 Vaccine

Ashiqur Rahman
ashiqur.r@niu.edu

Follow this and additional works at: https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations

Part of the Artificial Intelligence and Robotics Commons, and the Social Psychology Commons

Recommended Citation
https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations/7577

This Dissertation/Thesis is brought to you for free and open access by the Graduate Research & Artistry at Huskie Commons. It has been accepted for inclusion in Graduate Research Theses & Dissertations by an authorized administrator of Huskie Commons. For more information, please contact jschumacher@niu.edu.
ABSTRACT

MODELING AND VISUALIZATION OF LONG-TERM PUBLIC OPINION ON COVID-19 VACCINE

Ashiqur Rahman, MS
Department of Computer Science
Northern Illinois University, 2022
Hamed Alhoori, Director

The coronavirus pandemic created significant dependence on social media. While the social web was crucial in spreading timely information and informing the public, misinformation has also spread with little to no oversight. Several works have focused on identifying misinformation and topic analysis in COVID-19 (SARS-COV-2) tweets. While most of the previous studies focus on a shorter time frame, we analyzed a larger dataset starting from the beginning of the pandemic until the end of December 2021. Our work focuses on a novel area that identifies the motivating and demotivating topics of COVID-19 vaccination and analyzes these topics based on time, geographic location, and political orientation. We noticed that while the motivating topics mostly stay the same over time and geographic location, the demotivating topics vary rapidly. We developed an interactive visualization to understand the information better and find hidden relations between the public stance and the topics. Finally, we expanded our study to build a model to identify possible spamming by bots sharing scholarly articles. We found that health science literature is more spammed than other scientific areas.
NORTHERN ILLINOIS UNIVERSITY
DEKALB, ILLINOIS

DECEMBER 2022

MODELING AND VISUALIZATION OF LONG-TERM PUBLIC OPINION ON
COVID-19 VACCINE

BY

ASHIQUR RAHMAN
© 2022 Ashiqur Rahman

A THESIS SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
MASTER OF SCIENCE

DEPARTMENT OF COMPUTER SCIENCE

Thesis Director:
Hamed Alhoori
ACKNOWLEDGEMENTS

I have to thank so many people who helped me through my journey! First and foremost, I must acknowledge the support and confidence I received from my thesis supervisor – Dr. Hamed Alhoori. He trusted me to join his lab even though I had little research experience. He encouraged me and guided me to learn and improve myself and gave me the freedom to explore and discover my interests. I give my sincerest gratitude for his continuous support and guidance. I have to thank my committee members for their trust in me. Gratitude to Dr. Reva Freedman for guiding me with her knowledge. Her belief in me boosted my confidence to continue the work. Her remarks were essential for the refinement of the thesis. I am also grateful to Dr. Shupei Yuan for her interest and visions that helped me finalize this work and opened up several new possibilities for the future.

I am thankful to Dr. Ehsan Mohammadi at the University of South Carolina. His knowledge and suggestion were crucial to a significant part of my work.

I am also thankful for my colleagues in DATA Lab. They were always helping me with their knowledge and intellect throughout my journey. I can always go to them whenever I face a complex problem, and brainstorming with them is the best way to find a solution.

I am grateful to my friends, especially Iffat and Samiha, who encouraged me to return to academia. I wouldn’t be here without all their assistance and motivation.

Finally, gratitude to my wife for keeping up with me throughout this challenging path. Her support has been invaluable for getting this far.
DEDICATION

I dedicate this work to my mother, who devoted her life to a better future for her children. And my brothers, who are my motivation to do better.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>ix</td>
</tr>
<tr>
<td>List of Appendices</td>
<td>xi</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Dataset</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Objectives</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Methods</td>
<td>4</td>
</tr>
<tr>
<td>1.3.1 Machine Learning Models</td>
<td>4</td>
</tr>
<tr>
<td>1.3.1.1 BERTopic</td>
<td>4</td>
</tr>
<tr>
<td>1.3.1.2 DistilBERT</td>
<td>5</td>
</tr>
<tr>
<td>1.3.1.3 RoBERTa</td>
<td>5</td>
</tr>
<tr>
<td>1.3.1.4 Logistic Regression</td>
<td>5</td>
</tr>
<tr>
<td>1.3.1.5 Support Vector Model (SVM)</td>
<td>5</td>
</tr>
<tr>
<td>1.3.2 Libraries</td>
<td>6</td>
</tr>
<tr>
<td>1.3.2.1 GeoPy</td>
<td>6</td>
</tr>
<tr>
<td>1.3.2.2 D3</td>
<td>6</td>
</tr>
<tr>
<td>1.3.3 Performance Metrics</td>
<td>6</td>
</tr>
<tr>
<td>1.3.3.1 F-1 Score</td>
<td>7</td>
</tr>
<tr>
<td>1.3.3.2 Accuracy</td>
<td>7</td>
</tr>
<tr>
<td>Chapter</td>
<td>Section</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>1.3.3.3</td>
</tr>
<tr>
<td>2</td>
<td>Analyzing Twitter Data to Identify (De)Motivating Topics Related to COVID-19 Vaccine</td>
</tr>
<tr>
<td>2.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>2.1.1</td>
<td>Contributions</td>
</tr>
<tr>
<td>2.2</td>
<td>Related Works</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Misinformation and Vaccine</td>
</tr>
<tr>
<td>2.2.2</td>
<td>COVID-19 Misinformation Studies</td>
</tr>
<tr>
<td>2.2.3</td>
<td>COVID-19 Vaccine Sentiment and Stance Detection</td>
</tr>
<tr>
<td>2.2.4</td>
<td>(De)Motivation Studies and Vaccination Intent</td>
</tr>
<tr>
<td>2.3</td>
<td>Data Collection</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Twitter Dataset</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Training Dataset</td>
</tr>
<tr>
<td>2.3.2.1</td>
<td>COVID-19 Rumor Dataset</td>
</tr>
<tr>
<td>2.3.2.2</td>
<td>Avax Tweets Dataset</td>
</tr>
<tr>
<td>2.3.2.3</td>
<td>Authentic Vaccination Tweets Dataset</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Stance Ground Truth Dataset</td>
</tr>
<tr>
<td>2.4</td>
<td>(De)Motivating Topic Identification</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Classifying the Tweets</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Topic Analysis</td>
</tr>
<tr>
<td>2.4.2.1</td>
<td>Motivating Topics</td>
</tr>
<tr>
<td>2.4.2.2</td>
<td>Demotivating Topics</td>
</tr>
<tr>
<td>2.5</td>
<td>Stance Detection</td>
</tr>
<tr>
<td>2.6</td>
<td>Discussion</td>
</tr>
</tbody>
</table>
# Chapter 3: Visualizing Relation Between (De)Motivating Topics and Public Stance toward COVID-19 Vaccine

## 3.1 Introduction

## 3.2 Contributions

## 3.3 Related Works

## 3.4 Methods

### 3.4.1 Preparing the Datasets

### 3.4.2 Interactive Visualization Tool

#### 3.4.2.1 Visualizing Topics

#### 3.4.2.2 Visualizing Stance

#### 3.4.2.3 User Controls

## 3.5 Discussion

## 3.6 Future Work

## 3.7 Conclusion

---

# Chapter 4: Building a Model to Predict Possibility of Spamming by Twitter Bots on Scientific Articles

## 4.1 Introduction

### 4.1.1 Contributions

## 4.2 Related Works

### 4.2.1 Bot Activity in Social Media

### 4.2.2 Identifying Social Bots

### 4.2.3 Scholarly Publications on Social Media

### 4.2.4 Social Bots and Scholarly Articles
Chapter | Page
---|---
4.3 Methods | 56
  4.3.1 Data Collection | 56
  4.3.2 Data Processing | 57
  4.3.3 Building the Models | 63
  4.3.4 Statistical Analysis | 64
4.4 Discussion | 66
4.5 Conclusion | 67
5 Conclusion | 68
  5.1 Summary | 68
  5.2 Future Work | 69
References | 71
Appendices | 95
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Datasets, Sources, and Their Purpose</td>
<td>20</td>
</tr>
<tr>
<td>2.2 Models Trained to Identify (De)motivating Tweets</td>
<td>25</td>
</tr>
<tr>
<td>2.3 Topics from Motivating Tweets</td>
<td>28</td>
</tr>
<tr>
<td>2.4 Vaccine Topics from Motivating Tweets</td>
<td>29</td>
</tr>
<tr>
<td>2.5 Topics from Demotivating Tweets</td>
<td>31</td>
</tr>
<tr>
<td>2.6 Vaccine Topics from Demotivating Tweets</td>
<td>33</td>
</tr>
<tr>
<td>2.7 Transformer Training Performance</td>
<td>35</td>
</tr>
<tr>
<td>4.1 Metrics of Scores Provided by Botometer</td>
<td>57</td>
</tr>
<tr>
<td>4.2 Features Available in the Dataset</td>
<td>58</td>
</tr>
<tr>
<td>4.3 List of Features in the Twitter Dataset</td>
<td>59</td>
</tr>
<tr>
<td>4.4 Classification Report for Logistic Regression on the Twitter Dataset</td>
<td>64</td>
</tr>
<tr>
<td>4.5 Classification Report for Logistic Regression on the Health Dataset</td>
<td>64</td>
</tr>
<tr>
<td>4.6 Data Points for the Statistical Analysis</td>
<td>66</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Preparing the datasets with COVID-19-related tweets.</td>
<td>21</td>
</tr>
<tr>
<td>2.2</td>
<td>Number of tweets in different states (excluding Alaska, Hawaii, and distant territories).</td>
<td>21</td>
</tr>
<tr>
<td>2.3</td>
<td>Classification of tweets and identifying the topics and stance.</td>
<td>26</td>
</tr>
<tr>
<td>2.4</td>
<td>Hierarchy of topics related to “vaccine” in motivating tweets.</td>
<td>28</td>
</tr>
<tr>
<td>2.5</td>
<td>Top 10 topics in motivating tweets and their spread in the top 15 US states (by number of tweets).</td>
<td>30</td>
</tr>
<tr>
<td>2.6</td>
<td>Hierarchy of topics related to “vaccine” in demotivating tweets.</td>
<td>31</td>
</tr>
<tr>
<td>2.7</td>
<td>Top 10 topics in demotivating tweets and their spread in the top 15 US states (by number of tweets).</td>
<td>32</td>
</tr>
<tr>
<td>2.8</td>
<td>Stance toward vaccination (ignoring unrelated stance).</td>
<td>34</td>
</tr>
<tr>
<td>2.9</td>
<td>Vaccination stance in (de)motivating tweets in the top 10 states by number of tweets (ignoring unrelated stance).</td>
<td>35</td>
</tr>
<tr>
<td>2.10</td>
<td>Motivating tweets in top 10 states grouped by political affiliation in 2020 presidential election.</td>
<td>36</td>
</tr>
<tr>
<td>2.11</td>
<td>Demotivating tweets in top 10 states grouped by political affiliation in 2020 presidential election.</td>
<td>37</td>
</tr>
<tr>
<td>3.1</td>
<td>Screenshot of the visualization tool to explore topics and stance.</td>
<td>44</td>
</tr>
<tr>
<td>3.2</td>
<td>Topics visualized using bubble charts.</td>
<td>45</td>
</tr>
<tr>
<td>3.3</td>
<td>Cumulative stance over time visualized using a bubble chart.</td>
<td>46</td>
</tr>
<tr>
<td>3.4</td>
<td>User controls at the top-right.</td>
<td>47</td>
</tr>
<tr>
<td>3.5</td>
<td>Highlighting the tweets for one topic by mouse hover.</td>
<td>48</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.1</td>
<td>Altmetric score distribution before normalizing.</td>
<td>59</td>
</tr>
<tr>
<td>4.2</td>
<td>Correlation matrix of the dataset.</td>
<td>60</td>
</tr>
<tr>
<td>4.3</td>
<td>Spam activity by country (top 25).</td>
<td>61</td>
</tr>
<tr>
<td>4.4</td>
<td>Median bot scores by country.</td>
<td>61</td>
</tr>
<tr>
<td>4.5</td>
<td>Median bot scores by discipline.</td>
<td>62</td>
</tr>
<tr>
<td>4.6</td>
<td>ROC curve for logistic regression on the Twitter dataset.</td>
<td>64</td>
</tr>
<tr>
<td>4.7</td>
<td>ROC curve for logistic regression on the Health dataset.</td>
<td>65</td>
</tr>
<tr>
<td>4.8</td>
<td>Comparison of the ratio of spam activity on all research articles, health and human science areas, and other areas.</td>
<td>66</td>
</tr>
</tbody>
</table>
# LIST OF APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Appendix for Chapter 2</td>
<td>95</td>
</tr>
<tr>
<td>A.1 Datasets</td>
<td>96</td>
</tr>
<tr>
<td>A.1.1 Labeled Twitter Dataset</td>
<td>96</td>
</tr>
<tr>
<td>A.2 Machine Learning Models</td>
<td>96</td>
</tr>
<tr>
<td>A.2.1 (De)Motivation Classifier</td>
<td>96</td>
</tr>
<tr>
<td>A.2.2 Vaccine Stance Classifier</td>
<td>96</td>
</tr>
<tr>
<td>A.2.3 Topic Model</td>
<td>97</td>
</tr>
<tr>
<td>B Appendix for Chapter 3</td>
<td>98</td>
</tr>
<tr>
<td>B.1 Visualization</td>
<td>99</td>
</tr>
<tr>
<td>B.1.1 Visualization Tool</td>
<td>99</td>
</tr>
<tr>
<td>B.1.2 Source Code</td>
<td>99</td>
</tr>
<tr>
<td>C Appendix for Chapter 4</td>
<td>100</td>
</tr>
<tr>
<td>C.1 Datasets</td>
<td>101</td>
</tr>
<tr>
<td>C.1.1 Labeled Twitter Dataset</td>
<td>101</td>
</tr>
<tr>
<td>C.2 Machine Learning Model</td>
<td>101</td>
</tr>
<tr>
<td>C.2.1 Spam Prediction Model</td>
<td>101</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

The coronavirus pandemic has changed the way we live and work. The new world order introduced by this pandemic made us highly dependent on social media interactions and opinions, and the impact of this is profound. It has made us much more dependent on information disseminated on social media websites. This dependence can be positive or negative, depending on how well we use it to make informed decisions. We quickly share information that confirms our beliefs or biases, even with warnings of fake news [1].

This dependence on social media played a vital role in informing the public. However, this also opened up the floodgate of misinformation with very little oversight [2]. While vaccination is deemed the best option to beat the pandemic [3, 4], misinformation, among other issues, played an essential role in the low adoption rate [5, 6]. Therefore, it is crucial to identify the topics on social media that are (de)motivating the public about COVID-19 (SARS-COV-2) vaccination and address those topics accordingly.

This thesis addresses the abovementioned issues by analyzing COVID-19-related tweets spanning over two years, from January 2020 until December 2021. This thesis’s first contribution is identifying the topics that are (de)motivating about COVID-19 vaccination. We analyzed these topics to determine the transformation over time and in different US states and the political demographic of the states. We also analyzed the stance of the tweets and the relation between the topics and the stance. These topics can help the policymakers and healthcare organizations to focus their fight against the distrust of science and isolate the vulnerable groups to motivate them about COVID-19 vaccination. This process can be a steppingstone for a future crisis that requires mass public participation.
Secondly, we took a visualization approach to express the information gathered from the research. Inability to understand the vaccine development process and how the science works is one of the significant reasons for demotivation toward vaccination [7, 8]. We intended to visually explore Twitter data during this pandemic and find the most resonating topics among the general population that are (de)motivating them about the COVID-19 vaccination.

Finally, we intended to identify the use of scientific literature to mislead people. Previous research into Twitter data shows significant bot activity [9] that may steer the public opinion on crucial scientific discussions [10, 11, 12, 13]. In this part of the study, we built a model to identify the possibility of being spammed by Twitter bots for any academic article. This step leaves opportunities to expand and identify whether particular scientific articles are over-hyped by bots to steer public opinion and find any possible correlation between the (de)motivating topics and those articles. Additionally, this adds to the literature on understanding the societal impact of science [14, 15, 16].

1.1 Dataset

We used different sources to collect the data for the thesis. The first two sections used the Twitter dataset from Chen et al. [17], the COVID-19 rumor dataset from Cheng et al. [18], the “Avax Tweets” dataset – a COVID-19 vaccine hesitancy dataset from Muric et al. [19], and our collection of authentic tweets about COVID-19 vaccination. For the third section, we used the dataset from Altmetric.com [20] to collect the tweets and perform further analysis.
1.2 Objectives

While there are several smaller contributions of this research, in the bigger picture, the objectives are below:

1. Identify and analyze the (de)motivating topics about COVID-19 vaccination.
   - This analysis can identify how the topics differ based on geographic location, political orientation, and time.
   - Whether specific topics are resonating more within certain groups of people based on their stance toward vaccination.

2. Implement an interactive visualization to present the relation between (de)motivating topics and vaccination stance.
   - This can help us identify any hidden pattern and correlation between the vaccination stance and the topics.
   - Present an interactive and easy-to-understand method to visually represent scientific data to alleviate distrust toward science and better communicate a scientific message.

3. Identify the spamming by Twitter bots on scientific articles.
   - This can provide a better tool for interpreting scientific news on social media.
   - It opens up further research opportunities to identify bots’ possible use of scientific research to shape public opinion.
1.3 Methods

In their respective chapters, we have explained the methods used for different sections of this study. As an overview, for the first part, we used BERTopic [21] for topic modeling, a DistilBERT-based [22] model for tweet classification as (de)motivating, and a CT-BERT-based model to identify the stance of the tweets. We also considered RoBERTa-based [23] models for tweet classification and stance detection, but the models’ performances were unsatisfactory. To measure the performance of the models, we used accuracy and the Matthews correlation coefficient (MCC).

We used the data from the first part of the study and the D3 library [24] to build the visualizations for the second part.

Finally, for the third part of the study, we built a logistic regression model and a support vector model (SVM) to identify the spamming by bots. The logistic regression model performed better than SVM in our case. We used the F-1 score to measure the performance of the models in this case.

1.3.1 Machine Learning Models

1.3.1.1 BERTopic

BERTopic uses the Huggingface [25] transformers and c-TF-IDF to create clusters of interpretable topics. Using this model, it is also easier to see the progression of topics over time. c-TF-IDF is a variation of TF-IDF that takes the class into consideration [26].
1.3.1.2 **DistilBERT**

DistilBERT is a small and lightweight transformer model based on BERT that reduces the size of a BERT model by around 40%.

1.3.1.3 **RoBERTa**

RoBERTa is a variation of BERT with optimized pre-training method to achieve higher performance.

1.3.1.4 **Logistic Regression**

Logistic regression is a supervised learning algorithm used to predict the probability of a binary dependent variable.

1.3.1.5 **Support Vector Model (SVM)**

SVM is a supervised machine learning algorithm that tries to find a hyperplane in an N-dimensional space to classify the data points.
1.3.2 Libraries

1.3.2.1 GeoPy

GeoPy [27] is a Python library that works as a go-between to leverage the services from popular geocoding services.

1.3.2.2 D3

D3 is a JavaScript library for preparing interactive web-based visualizations.

1.3.3 Performance Metrics

Several different metrics are available to measure the performance of machine learning models. The following values are necessary to calculate the performance:

- True Positive (TP): The correctly predicted positive values.
- True Negative (TN): The correctly predicted negative values.
- False Positive (FP): The incorrectly predicted positive values.
- False Negative (FN): The incorrectly predicted negative values.

These four values together create the confusion matrix of any machine learning model.
1.3.3.1 **F-1 Score**

F-1 score is the harmonic mean of *precision* and *recall*. Here, precision is the proportion of true positives out of all the positive predictions – \( \frac{TP}{TP+FP} \). Furthermore, recall is the proportion of true positive predictions out of all the positive values – \( \frac{TP}{TP+FN} \).

\[
F-1 \text{ score} = \frac{2TP}{2TP+FP+FN} \tag{1.1}
\]

F-1 score considers \( TP \) as the value we are interested in and does not take \( TN \) into consideration.

1.3.3.2 **Accuracy**

Accuracy measures the ratio of correct predictions and total predictions. The measurement score ranges from 0 to 1, where 1 means 100% correct prediction.

\[
Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{1.2}
\]

1.3.3.3 **Matthews Correlation Coefficient (MCC)**

MCC considers all four values of the confusion matrix and works well even when the dataset is imbalanced. The value ranges from -1 to 1. The value of 1 represents a perfect prediction and -1 represents that the model always misclassifies. A value of 0 means the model’s classification is entirely random.

\[
MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{1.3}
\]
CHAPTER 2
ANALYZING TWITTER DATA TO IDENTIFY (DE)MOTIVATING TOPICS RELATED TO COVID-19 VACCINE

2.1 Introduction

Social media plays a vital role in modern life by availing us of communication, information dissemination, and steering social conversations [28]. These features have significantly more impact during a state of crisis [29]. We have seen this during major social events like mass shootings, natural disasters, national elections, and even anti-vaccination campaigns [10, 29, 30, 31].

The coronavirus pandemic created significant dependence on social media. While the social web was essential in spreading timely information and informing the public, misinformation has also spread with little to no oversight. As soon as the coronavirus emerged, racism, rumor, and fear-mongering started spreading like wildfire on different platforms [2]. Although there are projects such as Poynter [32] and EUvsDisinfo [33] actively monitoring and debunking false news, misinformation on social media is widely available. The World Health Organization (WHO) has partnered with major tech giants such as Facebook, Google, LinkedIn, Microsoft, Reddit, and Twitter to fight against misinformation [34]. However, misinformation is still widely available on these platforms. The WHO director-general called it the fight against “trolls and conspiracy theories.”

While vaccination has been proven to be the most effective tool for preventing diseases and keeping communities safe [3, 4], it will require a significant percentage of the popu-
lation to be vaccinated to achieve herd immunity [35] and leave the pandemic behind us. For comparison, measles requires 95% vaccination and polio requires 80% vaccination coverage to achieve this collective immunity [35]. Recent research suggests that although it might be unattainable to achieve herd immunity in the traditional sense against COVID-19 [36], a high vaccination rate will reduce the disease’s effect, make it more manageable [37], and return our life to normalcy. Although several vaccines have been developed and are accessible around the world [38], it has been proven difficult to achieve the necessary level of vaccination coverage worldwide. Different factors cause vaccine hesitancy, including but not limited to public trust in the development and approval of vaccines, economic disparity, education, and ethnicity [39, 40, 41, 42]. Misinformation on social media plays a vital role in emboldening misconceptions about vaccination. For instance, a survey in the UK revealed that people who relied on social web platforms to acquire information are more reluctant to receive vaccines [43]. Similarly, another study confirmed that the networks of anti-vaxxers were highly connected with doubtful clusters on Facebook [44], contributing to demotivation and distrust of the vaccination [45]. Politicizing the pandemic [5, 6] can also polarize the public and impact public trust in vaccination. Delay in achieving the vaccination target also impacts the economy. A recent study found that the total monetary harm due to “non-vaccination” in the United States is between 50 to 300 million dollars per day [46]. Therefore it is crucial to study, identify, and address the factors that are demotivating the public and increasing hesitancy and distrust toward vaccination.

In this chapter of our study, we primarily focused on the (de)motivation of getting vaccines during the COVID-19 pandemic and social media’s role in it. We aim to explore this issue on Twitter with the following research questions:
1. What are the most popular topics on Twitter that are (de)motivating people about the COVID-19 vaccine?

2. Which topics are influencing the public stance toward COVID-19 vaccination?

3. Do the motivating and demotivating topics about COVID-19 vaccine on Twitter change based on time, geographic location, or political landscape within the US?

2.1.1 Contributions

We answer the questions above throughout this chapter of the study and deliver the following data and machine learning models.

- A labeled Twitter dataset spanning from January 2020 to December 2021 containing, location, motivating status, vaccination stance, and topic label.

- An analysis of the COVID-19 vaccination topic distribution over time, US states, and political orientation of different states.

- A machine learning model to classify tweets as motivating or demotivating about the COVID-19 vaccine.

- A machine learning model to identify the COVID-19 vaccination stance of a user based on their tweets.

- Topic models to identify COVID-19 vaccination-related motivating and demotivating topics on Twitter.
2.2 Related Works

2.2.1 Misinformation and Vaccine

For better or worse, social media has become an increasingly popular method for everyday people to obtain information on various topics like scientific findings, current events, news, political occurrences, and many more. Social media can be an effective way for individuals to stay connected with the outside world and each other. Still, any user can post whatever content they desire, regardless of the validity of the information that the post contains. Thus, a paradox emerges in which everyday people have access to more information at their fingertips than ever before and an increased propensity for exposure to misinformation. This creates a chaotic information landscape characterized by a general inability of people to distinguish between fact and fiction in the pieces of information they encounter. This whirlwind of disseminated information and misinformation dramatically impacts the overall public perception of specific issues. Many researchers have attempted to analyze the relationship between social media trends and public opinion regarding public health issues like vaccination and immunization programs [47, 48, 49, 50, 51, 52, 53].

Several researchers have conducted analyses of Twitter content to determine the general public’s opinions on certain vaccines [10, 51, 52, 54, 55, 56]. For example, Becker et al. [51] have analyzed the contents of tweets (primarily posted by users in India, Indonesia, and Vietnam) containing sentiments regarding the pediatric pentavalent vaccine (DTP-HepB-Hib) [57] programs in those areas. They found that 37% of the tweets contained negative sentiments, while 63% were positive or neutral; they also indicated that most of the tweets contained links to websites or additional resources and did not add any
additional content or comments. Blankenship et al. [52] took this process a few steps further—tweets are not only analyzed for their sentiments about vaccines but also for their amount of engagement (retweets), categorizations of their content, and the types of curators that posted them. Results found no discernible variation in the number of times anti-vaccine tweets were retweeted across content categories. Twitter (12.9%), content curator “Trap It” (3.4%), and the Centers for Disease Control and Prevention (1.9%) were the top three domains among links in pro-vaccine tweets. Social media sites, including Twitter (14.9%), YouTube (8.4%), and Facebook (3.4%), were the most prevalent among the links in anti-vaccine tweets. The most frequently occurring theme in tweets with the hashtag #vaccineswork was childhood vaccination (40%). Vaccines could reduce outbreaks and deaths, according to 29% of tweets, which also referenced worldwide immunization efforts and improvement 21% of the time [52].

Similarly, other studies [10] have been conducted to determine which types of accounts are most problematic in spreading misinformation on Twitter. An analysis of sophisticated bots, Russian trolls, and content polluters found that regarding tweets about vaccines, Russian trolls can “amplify both sides” to create an online public discourse that can undermine public health; sophisticated bots, which are designed to look like legitimate accounts, can further undermine public health by increasing the number of those who hold apparent anti-vaccine sentiments.

Other studies [54, 55] explored whether users are more likely to post anti-HPV (human papilloma virus) vaccine tweets after being exposed to them themselves. Dunn et al. [54] found that the probability of tweeting something negative after being exposed to negative tweets was 37.78%, which was substantially higher than the likelihood of doing so after previously being exposed to neutral or positive tweets, which was 10.92%. In a subsequent study, Dunn et al. [55] expanded these results by attempting to create a model to explain the variance in HPV vaccine coverage. The study utilized an abundance of vari-
ables such as the exposure to HPV vaccine information on Twitter, socio-economic factors (e.g., poverty, education, insurance), racial and ethnic composition, and geographic location. They found that opinion exposure about the HPV vaccine on Twitter had more sway in determining vaccine coverage than socioeconomic factors.

Even more shocking, though, is the sheer volume of tweets containing information about vaccines—it has also been found that most of the tweets are posted by ordinary accounts (or lay consumers, i.e., not an academic, institution, or celebrity), and when sources are linked in tweets, it is also generally a link to a post made by an ordinary account [56]. It is generally agreed upon by these researchers [56] that Twitter can be an effective way to monitor opinions about public health issues and disseminate accurate information about the same issues. However, given the polarity and divisiveness of the current Twitter climate, and the sheer volume of tweets being sent out, Twitter itself and the overall information landscape must be improved (i.e., fact-checking, monitoring problematic accounts, improving overall information and media literacy standards, etc.) before that goal can become a reality.

Other researchers chose to explore this dynamic of diverse public opinion as it exists on a different platform – YouTube [53, 58, 59]. For instance, Basch et al. [53] viewed and categorized by poster 87 YouTube videos containing the phrases “vaccine safety” and “vaccines and children.” The three most common categories of video posters were ordinary consumers, internet or TV news, and individual health professionals; shockingly, 65.5% of the videos were deemed to be “anti-vaccine.” Similarly, after analyzing 172 YouTube videos related to the HPV vaccine for their tone and sentiment, response and reaction, and video source, Briones et al. [58] found that more than 51% of the YouTube videos contained negative HPV vaccination sentiments compared to 32% of positive ones, and the “anti-vaccine” videos are far more likely to be liked or shared than the positive videos. Conversely, another study [59] of HPV vaccination sentiments on YouTube found
that whether the video was positive or negative did not influence how many shares or views it received. The study also found that most videos could be classified as anti-vaccine.

Other researchers [60] have conducted similar analyses in other languages and geographic locations. For example, 123 Italian YouTube videos about vaccines were analyzed and the researchers discovered that 50% of the videos were positive in nature, 23% were negative, and 27% were neutral. Additionally, the study notes that both negative and positive videos alike utilized a “fear appeal” at a higher rate than any other persuasive strategy, like solidarity, economic interest, etc. YouTube videos posted regarding vaccines (both positive and negative) are rooted in fear and disdain for those with the opposite opinion. This further emphasizes the detrimental impact on public perception due to the dire state of online information seeking and sharing trends.

### 2.2.2 COVID-19 Misinformation Studies

COVID-19 is the first global pandemic in the social media era. This new experience opened up many nuances of social media and the fight against misinformation and fake news. Social media platforms have features such as automated bots that can facilitate the spread of misinformation [61]. Specifically, malicious activities have increased to an unprecedented level on social media during this pandemic [2]. The volume of COVID-19 misinformation led to dire consequences for the public and caused frontline workers to face even more challenges in stymieing the spread of the coronavirus. Public health agencies called this unchecked volume of mis/disinformation on social media platforms an infodemic [62, 63].
Researchers have been scrambling to keep up with the dissemination of misinformation. Islam et al. [62] collected articles from various online sources like fact-checking websites, social media, newspapers, and television networks to examine rumors, stigma, and conspiracy theories about COVID-19 and how they potentially impact individuals and communities. Lazer et al. [64] examined tweets by 1.6 million registered voters in the United States to determine who is sharing the misinformation and its sources. They determined that there is a strong political divide for sharing misinformation and is mostly shared by people over 50. They also found that the belief in misinformation is more prevalent in the younger population.

Evanega et al. [65] investigated the topics spreading misinformation during the early parts of the COVID-19 pandemic. They found that the majority of the misinformation was driven by “miracle cure” topic, and that prominent figures were the driving force in the spread of misinformation. They also noticed that only 16.4% of the overall conversation is about fact-checking or correcting the misinformation.

After analyzing 43.3 million tweets, Ferrara [13] found that automated social bots are used to disseminate COVID-19-related misinformation and political conspiracy theories. Al-Rakhami and Al-Amri [66] proposed a framework to use six different machine learning algorithms to detect COVID-19 misinformation. They collected the data using the Twitter API in the beginning of the pandemic and manually labeled the data to train the models.

Even vaccination to prevent COVID-19 is being debated, and misinformation is spread by the opponents of vaccination more frequently compared to the proponents [11]. Although officials are taking steps to handle misinformation regarding the vaccine [67], the efforts are still falling short [68] in tackling the diverse reasons [69] for the spread of misinformation.
Thelwall et al. [45] in their study found that while the majority of the vaccine hesitancy in the English language Twitter-sphere is related to right-wing conspiracy, there is a significant minority (18%) who are refusing the vaccine for non-political reasons like fear of being targeted as Black, development and approval speed, etc. Their study implies that vaccine hesitancy is not just confined to right-wing echo chamber, but can reach a wider audience.

2.2.3 COVID-19 Vaccine Sentiment and Stance Detection

Sentiment analysis is one of the major research areas in natural language processing (NLP) and can help us determine the overall perception of the population about any topic. Many researchers are doing sentiment analysis on tweets. Some of the sentiment analysis research during the COVID-19 pandemic shines a light on how people are responding to the pandemic [70, 71, 72].

Dubey [73] performed sentiment analysis on tweets from different countries between 11th and 31st March 2020. The researcher used the Syuzhet package [74] which classifies the tweets into eight different emotion categories. Within the data, Germany, France, the USA, and China showed balanced emotions between positive and negative tweets, whereas other countries showed a more positive attitude.

Manguri et al. [75] used the TextBlob Python library, a naïve Bayes sentiment classifier model, on tweets about COVID-19 for the week of 9th to 15th April 2020. The researchers found that people’s reactions vary from day to day and the majority of the tweets were neutral.

Stance detection is somewhat different than traditional sentiment analysis. While sentiment analysis can detect whether a text is positive, negative, or neutral, stance detection
can classify someone’s opinion as in favor or against a given target, which may or may not be present in the text, regardless of the emotion of the text [76]. Recent years have seen an increase in the use of social media metrics by researchers in order to measure the general public’s response to scientific results [77, 78, 79].

Augenstein et al. [80] worked on detecting stance toward a target topic that is not present in the tweet. They showed that conditional long short-term memory (LSTM) encoding is a suitable stance detection approach for an unseen target.

Dey et al. [81] proposed a two-phase support vector model (SVM) approach for stance detection on Twitter data. In the first phase, they classified the tweets into “neutral” and “other” (non-neutral). Then in the second phase, they classified the non-neutral tweets into “favor” vs. “against.” This method outperformed the state-of-the-art models.

Cotfas et al. [82] worked with tweets between 9 November 2020 and 8 December 2020, the month following the COVID-19 vaccine announcement, and found that the majority of tweets were in “neutral” territory and tweets in “favor” were greater than the “against” stance toward the vaccine.

Poddar et al. [83] extended the work of Cotfas et al. [82] by analyzing tweets from pre-COVID and post-COVID on data ranging from January 2018 to March 2021. They identified the stance of users toward the COVID-19 vaccine and analyzed the topics they are tweeting about to find a reason for the change in public stance.

2.2.4 (De)Motivation Studies and Vaccination Intent

While the majority of researchers consider misinformation to be the primary culprit for vaccine hesitancy and worked to identify it [40, 48, 62], there are many different factors like racial fear, stigma, economic constraints, and others, that can discourage peo-
ple from vaccination [39, 40, 41, 42]. Research suggests people are usually motivated by gain, altruism, or a protective attitude [7]. Protection motivation theory implies that severity and susceptibility increase vaccine intention [84].

Human psychology research proposes that there can be intrinsic motivation, where people are motivated by internal realization, and extrinsic motivation, where external forces steer people toward something [85]. And instead of competition, rewards, or threat of punishment, intrinsic motivation such as earning respect gains better results [86, 87].

Schmitz et al. [88] support the previous results that autonomous or intrinsic motivation works best for vaccine intention and uptake while controlled motivation (pressured by outside sources) does not work. They also noticed that people get more motivated by infection-related risk perception where personal health is at risk, rather than pandemic-related health concerns where overall societal health is considered. They also found that distrust toward science also impacts vaccine intention.

To acquire intrinsic motivation toward vaccination, understanding of vaccination and trust in the science are necessary. Lack of understanding of scientific findings, distrust toward politicians and the involvement of the federal government, fast-tracking and emergency authorization, concern about financial profits and political motives, and misrepresentation of the severity of COVID-19 are the primary reasons causing failure to motivate people toward vaccination [7, 8]. The efforts to motivate people and increase their vaccine knowledge fall short for several other reasons, including but not limited to unavailability of insurance reimbursement for consultation, lack of counseling, unavailability of vaccine during a clinic visit, and ease of getting an exemption [89].

Compared with previous studies, we analyzed Twitter data over a longer period of time that covers both before and after the rollout of major vaccines in the US and around the world. We also extracted the motivating and demotivating topics resonating in the Twitter-sphere regarding the COVID-19 vaccine and analyzed the spread of topics based
on geographic locations in the US. We then analyzed the public stance toward COVID-19 vaccine and identified the topics driving those stances. We also grouped the topics based on the political landscape of each state to investigate whether the misinformation tactics differ based on the majority political view of the area.

Our work in this study is novel in that we analyzed people’s stances over time to identify the (de)motivational topics influencing their stances toward the COVID-19 vaccine. While different programs and campaigns [90, 91] have been launched to educate people about COVID-19 and encourage the general public to vaccinate, our work can help identify the target parameters to have the most impact.

2.3 Data Collection

We have built a “Twitter Dataset” consisting of tweets and author information. Once we prepared the dataset, we used a machine learning classifier to classify the tweets as motivating or demotivating, identify the stance of the tweets, and extract the most prominent topics in both motivating and demotivating classes. We also prepared several smaller ground truth datasets to train the machine learning models. Table 2.1 lists the different datasets and their purposes.

2.3.1 Twitter Dataset

We collected close to 16 million tweets between January 2020 and December 2021 that contain information about COVID-19. We used the data from Chen et al. [17] by gathering the Tweet IDs from their GitHub repository [93]. Chen et al. used several keywords like “coronavirus”, “corona”, “COVID-19”, “pandemic”, “stayathome”, etc. to
Table 2.1: Datasets, Sources, and Their Purpose

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Source</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twitter dataset</td>
<td>Chen et al. [17]</td>
<td>Primary dataset containing the tweets and author information. We classified this data and performed analyses on it.</td>
</tr>
<tr>
<td>Training dataset</td>
<td>Cheng et al. [18], Muric et al. [19], Brandwatch [92]</td>
<td>Combination of three sources to build the ground truth dataset to classify (de)motivating tweets.</td>
</tr>
<tr>
<td>Stance ground truth dataset</td>
<td>Poddar et al. [83], Cotfas et al. [82], manual labeling</td>
<td>Combination of three labeled datasets to build the ground truth dataset for COVID-19 vaccine stance detection.</td>
</tr>
</tbody>
</table>

search for COVID-19-related tweets. In order to meet the rate limit of the Twitter API [94] and collect the data within a reasonable amount of time, we had to reduce the number of tweets. We randomly sampled at least 100,000 IDs each week and made sure that the data was stratified to match the distribution of the source dataset [93]. Then we used the Hydrator API tool [95] to collect all the tweet information. We collected 15,768,845 tweets using this method. Figure 2.1 shows the steps of the dataset creation.

After the data collection, we used the GeoPy [27] library to get the geographic location of the user from OpenStreetMap API [96]. In this step, we only considered the English-language tweets. Based on the location gathered using the API, we isolated the tweets from the United States and labeled each of the tweets by the respective US states and territories. We dropped the tweets without any geographic location for the authors, and for a few of the tweets we manually corrected any mislabeling of states with the help of other available information in each tweet such as zip code, landmark name, and so on. This is the primary dataset for the study. At the end of the process, we had 7,772,236 tweets in our dataset. Figure 2.2 shows the distribution of tweets by the US states.
Figure 2.1: Preparing the datasets with COVID-19-related tweets.

Figure 2.2: Number of tweets in different states (excluding Alaska, Hawaii, and distant territories).
2.3.2 Training Dataset

We created a dataset to train our machine learning models to classify the tweets in the Twitter dataset as motivating or demotivating. We combined data from three different sources to build a robust training dataset for our models. We combined the COVID-19 rumor dataset by Cheng et al. [18], the “Avax Tweets” dataset – a COVID-19 vaccine hesitancy dataset from Muric et al. [19], and our own collection of authentic tweets about COVID-19 vaccination.

2.3.2.1 COVID-19 Rumor Dataset

This dataset is a labeled dataset that contains COVID-19 rumors from both news sources and Twitter. The authors [18] of the dataset manually labeled 6,834 data points (4,129 rumors from news and 2,705 rumors from tweets). We used the text of the rumor and the label indicating whether the text was true, false, or unverified. We considered the “true” news and tweets as motivating for vaccination and the “unverified” and “false” as demotivating.

2.3.2.2 Avax Tweets Dataset

The authors [19] of the dataset curated a list of tweets that exhibit an antivaccine stance. The dataset contains over 1.8 million tweets over one year from October 2020 to November 2021. We have created a stratified sample of 100,000 tweets from the dataset and using the Hydrator API tool we collected 79,093 tweets from this sample. We consid-
ered these anti-vax tweets as demotivating tweets toward vaccination. We extracted the tweets from the dataset and labeled them as demotivating.

2.3.2.3 Authentic Vaccination Tweets Dataset

We collected historical tweets regarding COVID-19 vaccines from a curated list [97] of trusted sources from *Fortune* magazine. The list contained trusted public health officials, epidemiologists, virus experts, family doctors, and prominent health organizations. The authors of these accounts are sharing their experience in treating patients during the COVID-19 outbreak, sharing their advice, and refuting misinformation. This curated list gives us a source of authentic tweets regarding the pandemic and vaccination that are actively motivating people to vaccinate and advising the best ways to stay safe. We used the Brandwatch [92] API to collect COVID-19 vaccine-related tweets by the users in the aforementioned list between January 2020 and December 2022. We collected 19,992 tweets using this process, extracted the tweet texts, and labeled the tweets as motivating.

We merged the three datasets above to create the ground truth dataset for training our models. We labeled the dataset with a binary class indicating whether a tweet contains anti-vaccination rhetoric (demotivating) or not (motivating). This became our “Training dataset” containing two features – the text and the label. Before training the machine learning models, we cleaned the tweets by removing duplicates, retweet tags (“RT”), new-lines (“\n”), special characters, and words that contain non-English characters. Finally, we converted all the tweets to lowercase. After the cleanup, our dataset contained 60,647 demotivating tweets and 21,235 motivating tweets. We upsampled the motivating tweets to create a balanced dataset with 121,294 entries.
2.3.3 Stance Ground Truth Dataset

We extended the work of Poddar et al. [83] and Cotfas et al. [82] to identify the stance of tweets on the topic of vaccination. Although Poddar et al. [83] published their trained model, it did not perform well with newer tweets. After manually checking their results, we found that the stance prediction was correct for 55% and 61% for anti-vaccination and pro-vaccination respectively. We believe this was a result of the model being trained with data from a smaller timespan. Therefore, we decided to ignore their model and just use their labeled data in combination with our own and train a machine learning model with this newer data from a wider timespan. For this purpose, we manually labeled 1,064 tweets (469 in favor, 195 against, and 400 unrelated) and combined that with the data from Cotfas et al. [82] (991 in favor, 791 against, and 1,010 unrelated) and the data from Poddar et al. [83] (1,364 in favor, 490 against, and 1,285 unrelated). Finally, our ground truth data contained 6,995 entries with 2,824 in favor, 1,476 against, and 2,695 neutral tweets for the COVID-19 vaccine.

While manually labeling our own data, we used two annotators to make sure there was no bias in the labeling. The Cohen’s kappa score for the two annotators was 0.624, meaning the labeling of both annotators aligns at a satisfactory level. We also manually checked 100 random tweets from Cotfas et al. and 200 random tweets from Poddar et al. and our labeling aligned more than 80% of the time.
2.4 (De)Motivating Topic Identification

We used machine-learning models to classify the tweets from our Twitter dataset as motivating or demotivating. Then we used topic modeling to identify the prominent topics related to vaccines within the classified tweets. Figure 2.3 shows the steps of our analysis.

2.4.1 Classifying the Tweets

We fine-tuned DistilBERT [22] and RoBERTa [23], two different BERT-based pre-trained natural language processing (NLP) models, from the “Hugging Face transformers library” [98] and found DistilBERT to be the best. We tried different learning rates, batch sizes, and epochs to measure the performance. With a 70-30 train-test split, the accuracy for DistilBERT was 96.3%. Table 2.2 shows the performance of the training models we used.

Table 2.2: Models Trained to Identify (De)motivating Tweets

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>DistilBERT</td>
<td>96.30%</td>
</tr>
<tr>
<td>RoBERTa</td>
<td>73.40%</td>
</tr>
</tbody>
</table>

After training we used the model to classify the tweets in the Twitter dataset to identify the motivating and demotivating tweets. In this study, we created a small sample of 466,335 stratified tweets from the original dataset of 7,772,236 tweets. All the further analysis was done on this smaller sample.

The DistilBERT model classified 97,736 as motivating and 368,597 as demotivating tweets. After we classified the tweets, we used topic modeling to identify the vaccine-related topics from each of the classes.
Figure 2.3: Classification of tweets and identifying the topics and stance.

2.4.2 Topic Analysis

We used BERTopic [21] to create the topic models from our datasets. Although latent dirichlet allocation (LDA) [99] is one of the most popular algorithms to do topic
analysis, it takes some effort in hyperparameter tuning to generate meaningful topics and uses centroid-based topic extraction from document clusters. We also reviewed Top2Vec [100], a topic modeling algorithm that trains the document and word vectors jointly in a single semantic space. However, BERTopic leverages transformers [25] and c-TF-IDF to create dense clusters, allowing for easily interpretable topics. It is also possible to use pre-trained sentence transformer embedding models with BERTopic to find the prominence of any topic over time. In addition, the interactive visualization techniques make it much easier to investigate topic distribution using BERTopic.

We generated two separate topic models for the two classes in our dataset. After the training, each of the topic models returned the list of topics corresponding to the documents (tweets). We then extracted the top 20 topics related to “vaccine” from each of the topic models.

2.4.2.1 Motivating Topics

We fitted the model exclusively with the tweets classified as “motivating” to identify the motivating topics. The top 10 frequent topics from the model are displayed in Table 2.3 with the number of tweets in that topic and the top five words in that topic. We have ignored the most frequent topic that contains the stop words and pronouns.

Then we extracted “vaccine”-related topics from the topic model. Table 2.4 shows the top 20 “vaccine”-related topics from the motivating tweets. The score in the table represents the topic’s semantic similarity with the keyword (“vaccine”). The topic hierarchy in Figure 2.4 shows the relationship between topics in the set. Figure 2.5 shows the top 10 topics for each year and the frequency of tweets for each topic.
Table 2.3: Topics from Motivating Tweets

<table>
<thead>
<tr>
<th>Topic</th>
<th>Tweets</th>
<th>Top 5 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>0_mask_masks_wear_wearing</td>
<td>1452</td>
<td>mask, masks, wear, wearing</td>
</tr>
<tr>
<td>1_schools_students_school_learning</td>
<td>1290</td>
<td>schools, students, school, learning, teachers</td>
</tr>
<tr>
<td>2_amp_ov_optimistic_myself</td>
<td>684</td>
<td>amp, ov, optimistic, myself, ensure</td>
</tr>
<tr>
<td>3_football_game_players_basketball</td>
<td>480</td>
<td>football, game, players, basketball, games</td>
</tr>
<tr>
<td>4_sarscov2_sars_variant_infection</td>
<td>479</td>
<td>sarscov2, sars, variant, infection, genome</td>
</tr>
<tr>
<td>5_biden_joe_bidens_inaugural</td>
<td>435</td>
<td>biden, joe, bidens, inaugural, transition</td>
</tr>
<tr>
<td>6_music_album_song_spotify</td>
<td>431</td>
<td>music, album, song, spotify, songs</td>
</tr>
<tr>
<td>7_mail_voting_ballots_vote</td>
<td>423</td>
<td>mail, voting, ballots, vote, mailin</td>
</tr>
<tr>
<td>8_died_deaths_excess_1000</td>
<td>386</td>
<td>died, deaths, excess, 1000, per</td>
</tr>
<tr>
<td>9_corona_f<em>ck_sh</em>t_beer</td>
<td>351</td>
<td>corona, f<em>ck, sh</em>t, beer, b*tch</td>
</tr>
</tbody>
</table>

Figure 2.4: Hierarchy of topics related to “vaccine” in motivating tweets.
<table>
<thead>
<tr>
<th>Topic</th>
<th>Similarity</th>
<th>Top 5 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.desperately_vaccines_therapeutics_rollout</td>
<td>0.77255</td>
<td>desperately, vaccines, therapeutics, rollout, development</td>
</tr>
<tr>
<td>549_vaccineto_quicklyi_vaccinebut_bricker</td>
<td>0.64514</td>
<td>vaccineto, quicklyi, vaccinebut, bricker, coopting</td>
</tr>
<tr>
<td>89_shot_flu_never_fightflu</td>
<td>0.61227</td>
<td>shot, flu, never, fightflu, twindemic</td>
</tr>
<tr>
<td>231_misinformation_trust_pollquestion_trumpsvaccineisalie</td>
<td>0.60793</td>
<td>misinformation, trust, pollquestion, trumpsvaccineisalie, hash</td>
</tr>
<tr>
<td>248_slim_iscovid19_contraceptive_prevent</td>
<td>0.60710</td>
<td>slim, iscovid19, contraceptive, prevent, immune</td>
</tr>
<tr>
<td>1483_vaccinationshtha_adults_four_received</td>
<td>0.60689</td>
<td>vaccinationshtha, adults, four, received, leading</td>
</tr>
<tr>
<td>50_herd_immunity_natural_seroprevalence</td>
<td>0.58614</td>
<td>herd, immunity, natural, seroprevalence, tcell</td>
</tr>
<tr>
<td>33_flu_influenza_cold_season</td>
<td>0.57635</td>
<td>flu, influenza, cold, season, eradicated</td>
</tr>
<tr>
<td>1310_reinfection_suggesting_immunity_peop</td>
<td>0.55869</td>
<td>reinfection, suggesting, immunity, peop, hopef</td>
</tr>
<tr>
<td>531_vaccinat_protecting_stated_based</td>
<td>0.54879</td>
<td>vaccinat, protecting, stated, based, fully</td>
</tr>
<tr>
<td>200_antibodies_antibody_monoclonal_lotterythat</td>
<td>0.53910</td>
<td>antibodies, antibody, monoclonal, lotterythat, gsk</td>
</tr>
<tr>
<td>901_lungs_preventing_approved_vaccinations</td>
<td>0.51534</td>
<td>lungs, preventing, approved, vaccinations, normal</td>
</tr>
<tr>
<td>1554_scope_either_receiveReceived</td>
<td>0.50833</td>
<td>scope, either, receive, received, moderna</td>
</tr>
<tr>
<td>1463_sick1_sneezing_coughing_caring</td>
<td>0.50619</td>
<td>sick1, sneezing, coughing, caring, sanitizerwash</td>
</tr>
<tr>
<td>436_flu_strains_h3n2_longish</td>
<td>0.50240</td>
<td>flu, strains, h3n2, longish, paly</td>
</tr>
<tr>
<td>173_possiblebut_deferred_countless_kinesiology</td>
<td>0.49969</td>
<td>possiblebut, deferred, countless, kinesiology, covidbeing</td>
</tr>
<tr>
<td>1221_hiv_97_antibody_wow</td>
<td>0.49343</td>
<td>hiv, 97, antibody, wow, trials</td>
</tr>
<tr>
<td>91_molnupiravir_antiviral_clinical_drug</td>
<td>0.48830</td>
<td>molnupiravir, antiviral, clinical, drug, oral</td>
</tr>
<tr>
<td>544_administered_doses_cnn_ox</td>
<td>0.48290</td>
<td>administered, doses, cnn, ox, cnns</td>
</tr>
<tr>
<td>775_depre_commercials_rewarding_entitled</td>
<td>0.47337</td>
<td>depre, commercials, rewarding, entitled, 202021</td>
</tr>
</tbody>
</table>
2.4.2.2 Demotivating Topics

We also fitted the topic model with tweets classified as “demotivating.” Table 2.5 shows the top 10 most frequent topics from the model along with the number of tweets in each of the topics and the top five words in that topic. Pronouns and stop words are ignored as before.
### Table 2.5: Topics from Demotivating Tweets

<table>
<thead>
<tr>
<th>Topic</th>
<th>Tweets</th>
<th>Top 5 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.amp_wheelchairuser_haventwith_readthere</td>
<td>2723</td>
<td><em>amp, wheelchairuser, haven'twith, readthere, donaldyou</em></td>
</tr>
<tr>
<td>1.vaccinated_vaccine_vaccines_stacontrolling</td>
<td>1180</td>
<td><em>vaccinated, vaccine, vaccines, stacontrolling, rollout</em></td>
</tr>
<tr>
<td>2.cuomo_cuomos_andrew_nursing</td>
<td>1151</td>
<td><em>cuomo, cuomos, andrew, nursing, homes</em></td>
</tr>
<tr>
<td>3.nfl_reservecovid19_rogers_browns</td>
<td>871</td>
<td><em>nfl, reservecovid19, rogers, browns, aaron</em></td>
</tr>
<tr>
<td>4.nursing_homes_patients_elderly</td>
<td>801</td>
<td><em>nursing, homes, patients, elderly, governors</em></td>
</tr>
<tr>
<td>5.church_pastor_churches_worship</td>
<td>792</td>
<td><em>church, pastor, churches, worship, easter</em></td>
</tr>
<tr>
<td>6.cruise_ship_princess_navy</td>
<td>770</td>
<td><em>cruise, ship, princess, navy, passengers</em></td>
</tr>
<tr>
<td>7.vaccinated_vaccine_vaccines_minimizes</td>
<td>763</td>
<td><em>vaccinated, vaccine, vaccines, minimizes, downleadership</em></td>
</tr>
<tr>
<td>8.distancing_social_practicing_yip</td>
<td>738</td>
<td><em>distancing, social, practicing, yip, appa</em></td>
</tr>
<tr>
<td>9.corona_virusmeaning_enveloped_asteroid</td>
<td>709</td>
<td><em>corona, virusmeaning, enveloped, asteroid, dieme</em></td>
</tr>
</tbody>
</table>

**Figure 2.6:** Hierarchy of topics related to “vaccine” in demotivating tweets.
Similar to the earlier model, we extracted “vaccine”-related topics from the topic model. Table 2.6 shows the top 20 “vaccine”-related topics from the demotivating tweets.

![Figure 2.7: Top 10 topics in demotivating tweets and their spread in the top 15 US states (by number of tweets).](image)

The topic hierarchy in Figure 2.6 shows the relationship between topics in the set. Figure 2.7 shows the top 10 topics for each year and the frequency of tweets for each topic.
Table 2.6: Vaccine Topics from Demotivating Tweets

<table>
<thead>
<tr>
<th>Topic</th>
<th>Similarity</th>
<th>Top 5 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>1_vaccinated_vaccine_vaccines_stcontrolling</td>
<td>0.90244</td>
<td>vaccinated, vaccine, vaccines, sta-controlling, rollout</td>
</tr>
<tr>
<td>7_vaccinated_vaccine_vaccines_minimizes</td>
<td>0.87415</td>
<td>vaccinated, vaccine, vaccines, minimizes, downloadleadership</td>
</tr>
<tr>
<td>424_covidvaccine_vaccineside_effects_vaccinepasspos_fml</td>
<td>0.77769</td>
<td>covidvaccine, vaccinesideeffects, vaccinepasspos, fml, vaccinessaveslives</td>
</tr>
<tr>
<td>3642_immunity_natural_vaccinat_provides</td>
<td>0.66614</td>
<td>immunity, natural, vaccinat, provides, protection</td>
</tr>
<tr>
<td>369_immune_immunity_autoimmune_newslasting</td>
<td>0.65232</td>
<td>immune, immunity, autoimmune, newslasting, autoummune</td>
</tr>
<tr>
<td>922_getvaccinated_getvaccinated_now_resurrect_pandemichelp</td>
<td>0.63225</td>
<td>getvaccinated, getvaccinatednow, resurrect, pandemichelp, getvaccinatedasapthe</td>
</tr>
<tr>
<td>2243_immunocompromised_sacrifices_immunodeficient_guidelinescdc</td>
<td>0.58262</td>
<td>immunocompromised, sacrifices, immunodeficient, guidelinescdc, judgment</td>
</tr>
<tr>
<td>219_unvaccinated_992_995_peopleunvac</td>
<td>0.58117</td>
<td>unvaccinated, 992, 995, peopleunvac, foundunvaccinated</td>
</tr>
<tr>
<td>174_immunity_natural_itthat_isbig</td>
<td>0.57885</td>
<td>immunity, natural, itthat, isbig, covidnatural</td>
</tr>
<tr>
<td>4204_vaccinesaid_gonna_blacktwitter_cigarette</td>
<td>0.57702</td>
<td>vaccinesaid, gonna, blacktwitter, cigarette, yall</td>
</tr>
<tr>
<td>1166_vaccinesooo_testimony_getting_wowwatch_worldbecause</td>
<td>0.56898</td>
<td>vaccinesooo, testimonygetting, wowwatch, worldbecause, vaccinesbooster</td>
</tr>
<tr>
<td>2022_vaccinating_reopen_director_buoyed</td>
<td>0.56676</td>
<td>vaccinating, reopen, director, buoyed, teachers</td>
</tr>
<tr>
<td>6243_polio_gatesfunded_un_africa</td>
<td>0.56603</td>
<td>polio, gatesfunded, un, africa, 496000</td>
</tr>
<tr>
<td>41_biden_cancerwith_reordering_inherit</td>
<td>0.56335</td>
<td>biden, cancerwith, reordering, inherit, vaccin</td>
</tr>
<tr>
<td>139_unvaccinated_willfully_unnecc_onehundred</td>
<td>0.56173</td>
<td>unvaccinated, willfully, unnecc, onehundred, unvaccinatedbychoice</td>
</tr>
<tr>
<td>2423_patients_vaccinated_stro_favour</td>
<td>0.54486</td>
<td>heapatients, vaccinated, stro, favour, choicebecause</td>
</tr>
<tr>
<td>5660_vaccineortesting_courein_states_crossings_appeals</td>
<td>0.54226</td>
<td>vaccineortesting, courein-states, crossings, appeals, failedafghanistan</td>
</tr>
<tr>
<td>208_myst_covid19get_inconvenient_vaccinated</td>
<td>0.54093</td>
<td>myst, covid19get, inconvenient, vaccinatedif, fearcovidcovid</td>
</tr>
<tr>
<td>14_vaccinated_indoors_fully_wear</td>
<td>0.53886</td>
<td>vaccinated, indoors, fully, wear, distance</td>
</tr>
<tr>
<td>4278_bakker_vaccinate_jim_advice</td>
<td>0.53294</td>
<td>bakker, vaccinate, jim, advice, wash</td>
</tr>
</tbody>
</table>
2.5 Stance Detection

Figure 2.8: Stance toward vaccination (ignoring unrelated stance).

The stance of a text tells us whether the author is in favor or against a topic. To find the stance of users for the COVID-19 vaccine, we trained machine learning models using the labeled data discussed in Section 2.3.3 for stance detection. We used SimpleTransformer [101] to train models using Hugging Face [25] transformers. Table 2.7 shows the transformer training results reporting the performance in terms of the Matthews correlation coefficient (MCC). We used the MCC to measure performance because this provides a better understanding of the performance on an imbalanced dataset [102]. Since the ratio of pro-vaccine, anti-vaccine, and neutral are not balanced, the MCC provides better understanding than traditional accuracy or F1 score. The best model is highlighted in bold and referred to as ct-BERT-10 hereafter.
Table 2.7: Transformer Training Performance

<table>
<thead>
<tr>
<th>Transformer</th>
<th>MCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roberta (epochs: 10)</td>
<td>0.49</td>
</tr>
<tr>
<td>CT-BERT (epochs: 5)</td>
<td>0.58</td>
</tr>
<tr>
<td>CT-BERT (epochs: 10)</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Figure 2.9: Vaccination stance in (de)motivating tweets in the top 10 states by number of tweets (ignoring unrelated stance).

Then we classified the Twitter dataset using the ct-BERT-10 model to identify the stance of the tweets. After classifying the tweets, we found that the “Favor” stance surpasses “Against” in different time segments and US states. Figure 2.8 shows the number of tweets in each stance each month. Figure 2.9 shows the stance in different US states grouped by year and motivation classification. From Figure 2.8, we can see some interesting trends, like spikes in the “Favor” stance in the middle and the end of 2020.
2.6 Discussion

Figure 2.10: Motivating tweets in top 10 states grouped by political affiliation in 2020 presidential election.

The COVID-19 pandemic is a disrupting event changing our lives in more ways than we can imagine. After more than two years, we are still figuring out the impacts and learning to live with the new normal. This pandemic opened up new research horizons and questions about how we express ourselves on social media and how to handle a big challenge like this pandemic, motivate a large population toward a particular action, and prevent future disasters. Through our work, we noticed different interesting trends where public stance shifted over time and geographic location.

To answer our RQ1, we must first consider what motivation means in our context. The same statement can be either motivational or demotivational based on someone’s perspective. In our analysis in Section 2.4.2, we noticed that topics like face masks, schooling, and concern about deaths motivate people to get the vaccine. Sports news,
music, liberal politics, and presidential election news are also part of motivating topics. In contrast, the demotivating tweets contain topics like state-mandated vaccination, failures of liberal government, conservative politics, religion, and so on. There are very few overlaps like sports between the topics in the two different classes.

From Figure 2.4, we can see that there are two major branches in the motivating tweets. One segment is discussing policy, vaccine development and similar topics and the another is discussing symptoms and the dire physical impacts of the disease to motivate vaccination. Similarly, from Figure 2.6, we can see the three major branches of discussion that are demotivating – policy/politics, concern for the immunocompromised, and the push for vaccination. From the identified branches above, it is possible to target communities and encourage them toward vaccination and educate them about misconceptions.

With the stance detection analysis in Section 2.5, we explored our RQ2 and noticed that the tweets against COVID-19 vaccination contain more demotivating topics than their counterpart. There is also an intriguing trend that many “favorable” tweets include demo-
tivating topics, indicating that even well-intended tweets can have adverse effects. This trend may be due to political bias, religious beliefs, etc. We can expand this observation in future research.

We also noticed several spikes in positive stance toward the COVID-19 vaccine that warrant future exploration to identify any possible correlation with vaccine news, a new COVID-19 variant, or the political climate. Our study also exhibits future research possibilities for determining the exact topics influencing individual users’ stances.

While investigating RQ3, we noticed that the top motivating topics related to the COVID-19 vaccine stay mostly the same over time and geographic locations (Figure 2.5). On the other hand, the demotivating topics regarding the COVID-19 vaccine changed over time and are different based on geographic locations (Figure 2.7). We also noticed from Figure 2.10 and Figure 2.11 that the topics are different in liberal and conservative states during the same period. This observation opens up new avenues for research into the relationship between demotivating topics and regional politics.

2.7 Conclusion

Although several works have focused on identifying misinformation in the tweets about COVID-19 and performed topic analysis on them, there is little comprehensive work present to determine the tweeted topics that demotivate people from COVID-19 vaccination. Our work includes a large dataset ranging from the beginning of the pandemic until the end of 2021. We focused on a novel area that has the potential to identify the topics causing demotivation from COVID-19 vaccination. This work can be further extended to identify what drives people on social media platforms and can help address future social and public health issues.
CHAPTER 3
VISUALIZING RELATION BETWEEN (DE)MOTIVATING TOPICS AND PUBLIC STANCE TOWARD COVID-19 VACCINE

3.1 Introduction

Dependency on social media during the COVID-19 pandemic made the platforms essential for gathering information and forming public opinion. Misinformation and anti-vaccine propaganda also spread at an unprecedented level, making it difficult for healthcare organizations to reach their goals for vaccine coverage [2, 103]. It is essential to understand the information resonating on social media and playing an active role in (de)motivating the public [5, 6]. While public confidence in the scientific community is decreasing and trust in the vaccine development process is in question [7, 8], understanding the information producing these doubts can be critical to motivating people toward vaccination.

In this chapter of the study, we took a visualization\(^1\) approach to understanding the topics spreading on Twitter and (de)motivating the public. We used the dataset from Chapter 2 with some modifications and created an interactive visualization to explore the tweets and the topics. Through the visual exploration [104, 105], we noticed more demotivating topics than motivating ones, but the public stance is overall in favor of COVID-19 vaccination. One possible explanation is that, even when tweeting with good intentions, some language used in those tweets is backfiring and demotivating people from vaccina-

\(^1\)The visualization tool can be accessed from this URL: https://ashiqur-rony.github.io/visualize-covid-stance/
tion. Identifying these topics can help healthcare organizations steer public opinion and encourage the general public to vaccinate.

3.2 Contributions

Contributions of the study in this chapter are:

1. An interactive tool to visualize the topics and user stance with interactive features.

2. A visual exploration to understand the topics affecting public stance.

3. We also share our source code to help future researchers create a similar visual exploration tool.

3.3 Related Works

There is a thin line between fake news and misinformation. When verifiably false news is shared, that is called fake news [106]. On the other hand, misinformation is when false news is shared unintentionally [107, 108]. The mass of people can genuinely believe fake news without verifying it and thus take part in spreading misinformation.

After analyzing 43.3 million tweets, Ferrara [13] found that automated social bots are used to disseminate COVID-19-related misinformation and political conspiracy theories. Al-Rakhami and Al-Amri [66] proposed a framework using six machine learning algorithms to detect misinformation. They collected the data using the Twitter API at the beginning of the pandemic and manually labeled the data to train the models.

Misinformation is spread by the opponents of vaccination more frequently than by the proponents [11], who are widely debating the use of vaccination to prevent COVID-19.
Although officials are taking steps to handle misinformation regarding the vaccine [67], the efforts are still falling short [68] of tackling the diverse reasons [69] for the spread of misinformation. This seemingly unstoppable misinformation spread can change the population’s objectivity toward vaccination.

Researchers found that vaccine hesitancy goes hand-in-hand with believing in vaccine-related misinformation such as “cause people to catch COVID-19,” “more harmful than COVID-19,” and “will be used to alter people’s DNA” [40].

While the majority of research considered misinformation to be the primary culprit for vaccine hesitancy and worked to identify it [40, 48, 62], there are many different factors like racial fear, stigma, economic constraints, and many more that can discourage people from vaccination [39, 40, 41, 42]. Research suggests people are usually motivated by gain, altruism, or a protective attitude [7]. Protection motivation theory implies that severity and susceptibility increase the vaccine intention [84].

Sentiment analysis and stance detection can give a clear overview of the impact on public opinion. Sentiment analysis is one of the major research areas in natural language processing (NLP) that analyzes people’s attitudes and emotions from written language [109]. This can help us determine the overall perception of the population about any topic. Stance detection, however, is somewhat different from traditional sentiment analysis. While sentiment analysis can detect whether a block of text is positive, negative, or neutral, stance detection can classify someone’s opinion as in favor or against a given target, which may or may not be present in the text [76].

Several works on stance detection on Twitter data are available, including some related to COVID-19. Augenstein et al. [80] worked on detecting stances from tweets toward a target topic that is not present in the tweet. They showed that conditional long short-term memory (LSTM) encoding is a suitable stance detection approach for an unseen target.
Cotfas et al. [82] worked with tweets from the month following the COVID-19 vaccine announcement and found that the majority of tweets were in “neutral” territory and tweets in “favor” surpassed the “against” stance.

Dey et al. [81] proposed a two-phase support vector model (SVM) approach for stance detection on Twitter data. In the first phase, they classified the tweets into “neutral” and “other” (non-neutral). Then in the second phase, they classified the non-neutral tweets into “favor” vs. “against.” This method outperformed the state-of-the-art models.

Poddar et al. [83] extended the work of Cotfas et al. [82] by analyzing tweets from pre-COVID and post-COVID data ranging from January 2018 to March 2021. They identified the stance of users toward the COVID-19 vaccine and analyzed the topics they were tweeting to find a reason for the change in public stance.

Boon-Itt and Skunkan [110] analyzed tweets to find the topics and sentiments related to COVID-19. They found that the sentiment toward COVID-19 was mostly “negative,” although they did not try to relate the sentiment with topics.

Our study in this chapter visually represents the temporal change in people’s stance on COVID-19 vaccines and shows the (de)motivating topics impacting their opinion. There are several works on visualizing topics, including but not limited to word cloud, network, scatterplot, and heatmap [111]. “TopicPanorama” [112] and “Termite” [113] are two other interesting tools for visualizing topics.

Time series data visualization is also a widely researched area. Many different methods, like sequence chart, point chart, line chart, and bar chart, along with user interaction like zooming, brushing, and linking, present a variety of options for representing time series data [114].

Brehmer et al. [115] discussed the design choices for allowing temporal data to use a storytelling approach, which can better convey the message to users from different nar-
Wang et al. [116] proposed a visual analysis tool, “ConceptExplorer,” to analyze the concept drift using the time series data from different sources. Their tool lets users choose a timeline segment for analysis and compares concepts in a correlation matrix.

Carvalho et al. [117] chronicled the progress in COVID-19 research in the first 12 months from the outbreak of Sars-Cov-2. Their work summarizes the significant achievements of scientists and researchers in a short time while also outlining the gaps in research. They used a basic timeline visualization in their paper, which shows that a bare minimum visualization can be enough to convey the message.

In this chapter of our study, we used a simple bubble chart to represent the data because it is familiar to everyone and can convey several data points based on the bubbles position, size, and color.

### 3.4 Methods

#### 3.4.1 Preparing the Datasets

We used the classified and labeled dataset from Chapter 2 with some modifications to implement the visualizations. We filtered the data so that any user has at least 20 tweets in the whole dataset. Then we created a cumulative stance score for the tweets by adding the scores for each user over time. For this purpose, we considered the stance in favor of vaccination as +1, against as -1, and unrelated as 0, so that when someone is continuously in favor, their score increases over time and vice versa. We chose this for better exploration of public stance over time. If someone changes their stance, we can
quickly isolate them and examine their tweets. We also created a *topic frequency* feature containing the number of tweets on a topic each month.

Then we divided the dataset into two based on the (de)motivation classification of the tweets. This way, we can visualize the motivating and the demotivating tweets separately and examine the differences.

### 3.4.2 Interactive Visualization Tool

Figure 3.1: Screenshot of the visualization tool to explore topics and stance.
We developed a multiple-view visualization tool to explore the topics and stance from the COVID-19 vaccine-related datasets prepared above. Figure 3.1 shows a screenshot from the tool. We have explained each of the views present in the tool below.

### 3.4.2.1 Visualizing Topics

![Figure 3.2: Topics visualized using bubble charts.](image)

We used bubble charts to display the topics for a selected month. The horizontal axis represents the topics, while the vertical axis represents the topic frequency calculated...
in the dataset. The bubbles’ size represents the topic’s prominence at the selected time segment. We used different colors to isolate the topics in the visualization better.

We used animation to change the view when the user changes their selection (see User Controls). The hover effect also highlights the tweets in the stance visualization related to that topic. While visualizing the topics, we left the generic topic, which contained the stopwords and pronouns, out to avoid skewing.

3.4.2.2 Visualizing Stance

![Figure 3.3: Cumulative stance over time visualized using a bubble chart.](image)

We used bubble charts to visualize the cumulative stance over time as displayed in Figure 3.3. Each bubble in the visualization represents one tweet. The color of the bubble
differentiates the users. User selection of month can highlight the cumulative stance up
to that month, and the rest is grayed out. Users can also choose to highlight a single user
in the whole visualization (see User Controls).

We used animation to change the view when the user changes their selection. Hovering
the mouse pointer on a bubble shows details about that tweet, including the cumulative
stance score, location, primary topic, and the tweet text.

3.4.2.3 User Controls

![User Controls](image)

Figure 3.4: User controls at the top-right.

There are several user controls present in the visualization. A user can choose to
highlight a certain month, user, topic, or tweets. There are two drop-downs and a slider at
the top right of the visualization (or in the middle of two visualizations on mobile devices)
as displayed in Figure 3.4. Using these controls, the user can interact with and modify the
visualizations on the screen. The first drop-down lists the available data sources. The user
can select either motivating or demotivating tweets to populate the visualizations. From
the second drop-down, the user can select the author name to highlight tweets from that
author in the stance visualization. Using the range slider, the user can select the month
to display in the topic visualization and highlight the cumulative stance up to that month. All the controls have animation effects attached to them while changing the visualizations. There are also mouse controls attached to both topic and stance visualization. When the user hovers the mouse pointer on a topic, brushing techniques highlight that topic from the topic visualization at the top and highlight only the tweets using that topic from the stance visualization at the bottom, as displayed in Figure 3.5. The user can also make the selection from the other end. Hovering on a bubble from the stance visualization at the bottom highlights a single tweet, displaying the cumulative stance score, location, primary topic, and tweet text from that tweet text in the top-left corner. The primary topic of the selected tweet is also highlighted in the topic visualization at the top.

Figure 3.5: Highlighting the tweets for one topic by mouse hover.
3.5 Discussion

By analyzing (de)motivating topics on Twitter and user stances regarding COVID-19 vaccine, we noticed that demotivating tweets are more frequent than motivating ones. We also strangely noticed that most demotivating tweets had a favorable stance toward vaccination. This observation led us to believe that while tweeting with good intentions, political polarization, the confirmation bubble, or other factors are demotivating the public about vaccination. This finding opens up new research areas to explore.

We also noticed that religion and politics are prominent in the demotivating tweets, while schools, education, and statistical facts are more evident in the motivating tweets. The prominent topics also changed after December 2020, which warrants further study to understand whether the election year played a vital role in shaping public opinion.

3.6 Future Work

There are several possible improvements to the study. First of all, we only highlighted the primary topics from the tweets. We can use the presence of alternative topics in the tweets to understand the niches better.

Secondly, we only display the name of the state when the user hovers their mouse on a tweet. Implementing a feature to isolate tweets based on the state will give better exploration tools and can let us discover unforeseen facts.

Finally, the visualization works best on a larger screen. However, with the popularity of mobile devices and information consumption on smaller screens, a mobile-friendly version will benefit researchers and the general public.
3.7 Conclusion

In conclusion, we believe this is an important research area that merits future expansion to understand and tackle (de)motivation and public understanding of science to restore public trust in the scientific method. The methods derived from this study can also be extended for any future crisis to tackle the misinformation campaign.
CHAPTER 4
BUILDING A MODEL TO PREDICT POSSIBILITY OF SPAMMING BY TWITTER BOTS ON SCIENTIFIC ARTICLES

4.1 Introduction

Dissemination of academic articles through social media can be a valuable metric to determine their impact on the public. Altmetric gives us precisely that as a numeric score for any academic article. Altmetrics, or alternative metrics [120, 121, 122, 123, 124], monitor academic mentions on online platforms in order to study and quantify the influence of scholarly products. Altmetrics have been used to predict scholarly citations [125, 126], mentions in news [127], public policy documents [128, 129], patents [130], and long-term online mentions [131]. It includes peer reviews on the Faculty of 1000, citations on Wikipedia, discussions on research blogs, mainstream media coverage, bookmarks on reference managers like Mendeley, and mentions on social networks such as Twitter [20]. Because of its high reliance on social media activity, the Atlmetric score is vulnerable to bot activities on social media.

In our current study, we only look at the Twitter dataset from Altmetric and the activity of Twitter bots in that dataset. The term “bot” here refers to automated social media accounts that act with specific goals and prevent the organic growth of a conversation on

---

1Part of this chapter was published in Social Media & Society, 12th Intl. Conf., as a work-in-progress paper [118].

The complete text of this chapter is currently under review by the web conference [119] 2023 as collaborative work with co-authors Dr. Hamed Alhoori of Northern Illinois University and Dr. Ehsan Mohammodi of the University of South Carolina.
the social media platform [132, 133]. Besides the Twitter user information and Altmetric score, the dataset contains several important features for each article, like the paper’s origin, publication date, research area, and where it was published.

Being one of the most prominent social media globally, with 126 million daily active users and about 330 million monthly users [134], Twitter has become a preferred platform for scholarly communication. However, Twitter is also prone to bot and spam activity. Spambot accounts on this platform are especially problematic because they can quickly generate vast traffic and influence people’s opinions [9]. Much research on Twitter data focuses on identifying bots [135, 136, 137, 138], and there is a growing concern that bot activity on social media may significantly change the public opinion on crucial scientific discussion [10, 11, 12, 13].

With the increasing influence of social media on everyday life, many research areas and researchers may intentionally or unintentionally become targets of social bots [12, 139, 140]. Therefore, in our research, we use features of articles to predict the possibility of being spammed using different machine learning models.

Although Twitter bot activity is a well-researched topic, our research is directed to a novel approach to building a machine learning model to predict the possibility of whether Twitter bots will spam any given research article. This model can be helpful for researchers, decision makers, and the general population as a whole. We also identified that health and human science articles are more prone to spamming by bots than other research areas.

We discuss related works in the next section, followed by data collection methods, data processing, building models, and statistical analysis. We complete this chapter of the study with a discussion of our findings and possible future expansions of the work.
4.1.1 Contributions

• A comprehensive dataset of scholarly articles with binary labels indicating possible spamming by Twitter bots.

• A machine learning model to identify possible bot activity on an academic article.

• An analysis to determine the probability of increased spamming by bots in the health and human science area.

4.2 Related Works

4.2.1 Bot Activity in Social Media

Social bots are a long-studied but unsolved problem in the research community [141]. The frequency of social bots and their activity significantly thwarts different scientific advancements. Besides creating a nuisance and wasting valuable time, they can shift public opinion [10] and have a real-world impact. We see that with the anti-vaccination movement [10, 139], spreading conspiracy theories [11, 13, 62, 142], and most recently with COVID-19-related misinformation dissemination [11, 13]. Besides political and nefarious purposes, businesses also use bots to boost their sales and inflate their presence on social media [143, 144].

Although not all bots are malicious [138], some have specific agendas and can cause severe social discord [10, 142, 145]. Researchers have tried to understand the impact of bot activity and how it may steer social media conversation. Monsted et al. [146] examined whether information adoption is done with simple exposure (single source) or
complex exposure (multiple sources). The authors used social bots (39 coordinated bots) to spread information and determine the effect. They conclude that the complex contagion model spreads information more reliably.

4.2.2 Identifying Social Bots

Bot activity and detection on social media have been a significant area of research interest. Bot activity makes it difficult to build any actual metric or find reliable results by analyzing social media data to make an informed decision. Ongoing research proposes different methods for identifying and cleaning bot activity from social media data. Minnich [147] proposed removing all noise from the text using NLP techniques, “behavior profiling,” and “BotWalk” [136] to remove bot activity from the data. For “Behavior Profiling,” an extensive collection of features for any user is compiled and analyzed to determine bot-like behavior. Then the “BotWalk” algorithm uses the behavioral feature vector to identify potential bots in real time.

Bots frequently change their approaches to social media, making it very difficult to detect their behavior [148]. This adaptive behavior makes improving the bot detection algorithms a continuous process. Several methods to detect bots and spammers have been proposed [9, 135, 136, 143, 149, 150, 151, 152, 153] by researchers over the years. Chavoshi et al. [135] proposed the bot detection method “DeBot,” which uses warped correlation. The proposed method detects bots based on synchronized and correlated activity between users. Another approach that Chavoshi et al. [137] proposed depends on correlated activities between Twitter accounts. They used Amazon Mechanical Turk to validate the detections. Minnich et al. [136] proposed another near real-time bot detection method. Since modern bots keep adapting their behavior to evade detection, the proposed
method – BotWalk – uses a model that evaluates different Twitter account features to identify possible bot accounts. Efthimion et al. [9] analyzed the prevalence of bots and introduced a new algorithm to detect Twitter bots, achieving a low of 2.25% misclassification rate. Feng et al. [154] proposed a comprehensive dataset to benchmark different bot detection methods and their adaptability with the evolution of bot behavior.

Martini et al. [155] analyzed three different bot detection techniques and concluded that different methods provide vastly different results, causing a reliability problem in bot detection tools.

Considering all these, we used Botometer [156] to identify bots because this tool gives us detailed scores in different metrics and gives us the control to set the threshold to label a Twitter user as a bot.

### 4.2.3 Scholarly Publications on Social Media

With the high popularity of social media, more research is being shared and discussed on these platforms. However, the popularity and frequency of shares on social media do not always translate into citations. Didegah and Thelwall [157] conclude that counts of tweets are not a reliable metric because many researchers may tweet or save articles from other researchers they follow or work with but are not quite interested in the area. The actual correlation between tweets and citations was very low. Bot activity also skewed the measurement. The experiment found that among all the Twitter users in their dataset, almost 46% of prolific article tweeters were bots while 21% of moderate and 11% of occasional article tweeters were bots.
4.2.4 Social Bots and Scholarly Articles

Leidl [158] tried to find a correlation between the broader dissemination of research papers by bots and citations of those papers but could not find any conclusive relation between those. Ortega [159] showed that the journals with their own Twitter accounts get more tweets and citations than those without. Haustein et al. [160] found that automated Twitter accounts create a significant number of tweets to scientific articles.

Even though the impact of social media activity on academic achievement is inconclusive [12, 161, 162], social media is still a powerful platform for bridging the communication gap between researchers and the general public. Even though immediate engagement and feedback encourage sharing on social media [163, 164, 165], this can give a false perception about public interest. Massive amounts of bot activity can also play an important role and skew public interpretation of scientific research — a better understanding of why and which research areas the bots target is essential to prevent that.

4.3 Methods

4.3.1 Data Collection

We used data from Altmetric [20], Twitter API [94], and Botometer [156]. Altmetric provides alternative metrics based on online activity for academic articles, including social media reach, blog posts, news mentions, videos, and forums. We used the Twitter API to collect information about the user who posted the article on Twitter. Finally, the Botometer API gave us a bot score for each Twitter handle.
The dataset from Altmetric contained more than 10 million records. We found 182,277 unique Twitter accounts that posted these 10 million articles. We collected information about all these Twitter accounts using the Twitter API. Finally, we used the Botometer API to assign a bot score to each user account.

4.3.2 Data Processing

The Botometer API provides a bot score for any given Twitter account based on several metrics such as their current activity, the historical tweets of the user, and profile information. We get a score on a scale of 0 to 5 for the metrics displayed in Table 4.1.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td>User’s tweet content score</td>
</tr>
<tr>
<td>English</td>
<td>Whether the content is in English</td>
</tr>
<tr>
<td>Friend</td>
<td>How many friends/followers the user has</td>
</tr>
<tr>
<td>Network</td>
<td>Friend network of the user</td>
</tr>
<tr>
<td>Sentiment</td>
<td>Sentiment of the user’s tweets</td>
</tr>
<tr>
<td>Temporal</td>
<td>User’s behavior over time</td>
</tr>
<tr>
<td>Universal</td>
<td>An overall score for the user</td>
</tr>
<tr>
<td>User</td>
<td>User’s profile information</td>
</tr>
</tbody>
</table>

We calculated a bot score between 0 and 40 by summing up the scores for each Twitter account and adding this as a new feature to our dataset. Then we aggregated the dataset into groups of unique articles and gave a bot score based on the median score of all the tweets in each group. The median value was chosen instead of the mean to prevent the scores from being biased by outliers. This aggregation resulted in around 1.4 million records with an overall bot activity score for each scholarly article. After this process, we had the features listed in Table 4.2.
Table 4.2: Features Available in the Dataset

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altmetric ID</td>
<td>Unique ID for each entry</td>
</tr>
<tr>
<td>Scopus</td>
<td>Research area of the paper</td>
</tr>
<tr>
<td>Twitter poster types</td>
<td>An indication of the tweet author (e.g., researcher, science communicator, public, etc.)</td>
</tr>
<tr>
<td>Paper pubdate</td>
<td>Publication date of the paper</td>
</tr>
<tr>
<td>First seen on</td>
<td>First time seen on social media</td>
</tr>
<tr>
<td>Last mentioned on</td>
<td>Last time seen on social media</td>
</tr>
<tr>
<td>Subjects</td>
<td>Subjects covered in the paper (sub-groups of “scopus” above)</td>
</tr>
<tr>
<td>Selected quotes</td>
<td>Text quoted in the tweet along with the paper’s link</td>
</tr>
<tr>
<td>Funders</td>
<td>Funders of the paper</td>
</tr>
<tr>
<td>Twitter unique users</td>
<td>Number of unique Twitter users sharing the paper</td>
</tr>
<tr>
<td>Twitter posts count</td>
<td>Number of tweets sharing the paper</td>
</tr>
<tr>
<td>Journal</td>
<td>Journal that published the paper</td>
</tr>
<tr>
<td>Research type</td>
<td>Whether the paper is an article or news</td>
</tr>
<tr>
<td>Publisher</td>
<td>Publishing company of the research</td>
</tr>
<tr>
<td>Altmetric score</td>
<td>Altmetric attention score for the paper based on all online activity</td>
</tr>
<tr>
<td>Authors</td>
<td>List of authors of the paper</td>
</tr>
<tr>
<td>Tweet ID</td>
<td>Unique tweet ID from the Twitter API</td>
</tr>
<tr>
<td>Twitter desc</td>
<td>Profile description of the Twitter user</td>
</tr>
<tr>
<td>Twitter ID</td>
<td>Unique user ID of the Twitter user</td>
</tr>
<tr>
<td>Twitter author followers</td>
<td>Number of followers of the Twitter user</td>
</tr>
<tr>
<td>Twitter author name</td>
<td>Name of the Twitter user</td>
</tr>
<tr>
<td>Author loc</td>
<td>Geographic location of the Twitter user</td>
</tr>
<tr>
<td>Tweet posted on</td>
<td>Date of the tweet</td>
</tr>
<tr>
<td>Retweeters</td>
<td>Twitter users who retweet the original tweet</td>
</tr>
<tr>
<td>Author ID</td>
<td>Twitter handle of the user</td>
</tr>
<tr>
<td>Overall score</td>
<td>Overall Botometer score for the article</td>
</tr>
</tbody>
</table>

However, many of the features are related, like “scopus” and “subjects,” “journal” and “publisher,” and so on. We removed the redundant features, preferring the broader feature that covered more generalized information. For example, “scopus” can be “Medicine,” and the “subjects” can be “Orthopedic,” and we chose to keep the “scopus” feature.

Similarly, all the Twitter user features like follower count, user description, etc., are taken into account by Botometer, so we removed the unnecessary features before building the model.
Based on the bot score, we created a new binary feature called “is_Spammed,” which is set to true when the overall bot score is above 20. This threshold of 20 gives us 201,679 records flagged as spammed and 1,398,007 as not spammed. We noticed that the Altmetric score is vastly imbalanced, ranging from 0.25 to 8,268.56, as displayed in Figure 4.1. We normalized the score to a value between 0 and 1. At the end of these steps, our final dataset, the Twitter dataset, had seven features, as listed in Table 4.3.

![Figure 4.1: Altmetric score distribution before normalizing.](image)

Table 4.3: List of Features in the Twitter Dataset

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scopus</td>
<td>Research area of the paper</td>
</tr>
<tr>
<td>Journal</td>
<td>Journal that published the paper</td>
</tr>
<tr>
<td>Research type</td>
<td>Whether the paper is an article or news</td>
</tr>
<tr>
<td>Publisher</td>
<td>Publishing company of the research</td>
</tr>
<tr>
<td>Altmetric score</td>
<td>Altmetric attention score for the paper based on all online activity</td>
</tr>
<tr>
<td>Author location</td>
<td>Geographic location of the Twitter user</td>
</tr>
<tr>
<td>Is Spammed</td>
<td>Binary feature to determine whether an article is spammed</td>
</tr>
</tbody>
</table>
A correlation matrix (Figure 4.2) of the remaining features gives a clear picture of the dataset.

![Correlation Matrix](image)

**Figure 4.2:** Correlation matrix of the dataset.

We noticed distinctive trends in the spam activity while analyzing the dataset. By grouping our data based on user location, we observed a high spamming rate in the USA, UK, and Spain, as shown in Figure 4.3. We noticed that the median bot score is relatively high in some countries while others are moderately below our threshold of 20, as displayed in Figure 4.4. Similarly, grouping the articles by research area, we notice that the median bot score is highest in “Immunology and Microbiology,” followed by “Energy” and “Pharmacy, Toxicity, and Pharmaceutics.”

From our analysis, we noticed that papers from health and human science disciplines are more frequent and consequently have relatively more bot activity, as shown in Figure
Figure 4.3: Spam activity by country (top 25).

Figure 4.4: Median bot scores by country.
Figure 4.5: Median bot scores by discipline.

4.5. We isolated the articles for those disciplines by filtering the Twitter dataset and considering only the “scopus” features listed below.

- Biochemistry
- Genetics and Molecular Biology
- Medicine
- Life Sciences
- Health Sciences
- Psychology
- Dentistry
• Health Professions

• Nursing

• Pharmacology, Toxicology, and Pharmaceutics

• Immunology and Microbiology

• Neuroscience

This filtering resulted in 1.17 million records out of the original 1.4 million. We conducted further training and analysis on this dataset. We’ll refer to this filtered data as the “Health dataset.”

4.3.3 Building the Models

With the bot score threshold of 20 for spam, we had around 15% of the data marked as spammed. We upsampled the spammed entries using Scikit-Learn’s resample method to create a balanced dataset. Then we built a classifier model to predict an article’s spamming probability. We built a logistic regression model for the dataset with a 70-30 train-test split. The logistic regression on the dataset got an F-1 score of 0.70. The classification matrix for the model is displayed in Table 4.4 and the ROC curve for the model is displayed in Figure 4.6.

Then, we used the same logistic regression model on the Health dataset and had an F-1 score of 0.70. Table 4.5 and Figure 4.7 show this model’s classification report and ROC curve, respectively.

The performance is almost the same on the Health dataset and the Twitter dataset, which is expected as the Health dataset is, in fact, a majority subset of the Twitter dataset.
Table 4.4: Classification Report for Logistic Regression on the Twitter Dataset

<table>
<thead>
<tr>
<th></th>
<th>precision</th>
<th>recall</th>
<th>F1-score</th>
<th>support</th>
</tr>
</thead>
<tbody>
<tr>
<td>False</td>
<td>0.69</td>
<td>0.72</td>
<td>0.70</td>
<td>359185</td>
</tr>
<tr>
<td>True</td>
<td>0.71</td>
<td>0.67</td>
<td>0.69</td>
<td>358612</td>
</tr>
<tr>
<td>accuracy</td>
<td>0.70</td>
<td></td>
<td></td>
<td>717797</td>
</tr>
<tr>
<td>macro avg</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
<td>717797</td>
</tr>
<tr>
<td>weighted avg</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
<td>717797</td>
</tr>
</tbody>
</table>

Figure 4.6: ROC curve for logistic regression on the Twitter dataset.

Table 4.5: Classification Report for Logistic Regression on the Health Dataset

<table>
<thead>
<tr>
<th></th>
<th>precision</th>
<th>recall</th>
<th>F1-score</th>
<th>support</th>
</tr>
</thead>
<tbody>
<tr>
<td>False</td>
<td>0.69</td>
<td>0.72</td>
<td>0.71</td>
<td>301728</td>
</tr>
<tr>
<td>True</td>
<td>0.71</td>
<td>0.67</td>
<td>0.69</td>
<td>300198</td>
</tr>
<tr>
<td>accuracy</td>
<td>0.70</td>
<td></td>
<td></td>
<td>601926</td>
</tr>
<tr>
<td>macro avg</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
<td>601926</td>
</tr>
<tr>
<td>weighted avg</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
<td>601926</td>
</tr>
</tbody>
</table>

4.3.4 Statistical Analysis

We observed from basic plots that health and human science articles are more likely to be spammed. We performed a hypothesis test to determine whether the higher level of spamming is statistically significant. For this purpose, we considered the following
Figure 4.7: ROC curve for logistic regression on the Health dataset.

**hypothesis:**

**Null hypothesis:**

Spam activity in health and human science compared to other research disciplines is not significantly higher.

**Alternate hypothesis:**

Health and human science articles have more spam activity than the other research disciplines.

If we plot the ratio of spammed articles, as displayed in Figure 4.8, we can see that the health and human science discipline has slightly higher spam activity.

We performed a two-tailed Z-test to validate our hypothesis. We considered a confidence level of 99% required to reject the null hypothesis, so the p-value must be smaller than 0.01. Table 4.6 shows all the data points for the analysis.

Once we performed the Z-test, we got a p-value of $2.12 \times 10^{-232}$, close to zero, and a Z-score of 32.5 standard deviations to the right of the center. Based on this test, we can
Figure 4.8: Comparison of the ratio of spam activity on all research articles, health and human science areas, and other areas.

Table 4.6: Data Points for the Statistical Analysis

<table>
<thead>
<tr>
<th>Data point</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All research areas</td>
<td></td>
</tr>
<tr>
<td>Number of articles</td>
<td>1,398,007</td>
</tr>
<tr>
<td>Articles flagged as spam</td>
<td>201,679</td>
</tr>
<tr>
<td>Health &amp; human science areas</td>
<td></td>
</tr>
<tr>
<td>Number of articles</td>
<td>1,178,085</td>
</tr>
<tr>
<td>Articles flagged as spam</td>
<td>174,876</td>
</tr>
<tr>
<td>Other areas</td>
<td></td>
</tr>
<tr>
<td>Number of articles</td>
<td>219,922</td>
</tr>
<tr>
<td>Articles flagged as spam</td>
<td>26,803</td>
</tr>
</tbody>
</table>

reject the null hypothesis and conclude that health and human science articles have more spam activity than other research areas.

4.4 Discussion

This study gives us a critical perspective on disseminating scholarly articles on social media. While there is increasing distrust toward science and lack of understanding of scientific method in society for different reasons [166, 167, 168], the model we developed can validate the popularity of any given scholarly article on social media. This model does not argue the authenticity or validity of the actual work; instead, it can only predict whether the dissemination of the article is achieved by bot activity. This prediction can
be helpful for both researchers and policymakers to understand public acceptance of any given scientific publication.

Our study also concluded that bots targeted the health and human science discipline more. Although we could not identify a definitive reason for that, we hypothesize that since the general public is directly connected with health and human science, and they care about diseases, vaccination, and the health care system, topics in this area in turn attract more bots and spamming. We believe this can be an exciting idea for future research.

4.5 Conclusion

In this chapter of the study, we examined how Altmetrics features combined with the information about Twitter users can predict if a scholarly paper will attract bots on Twitter. Previous work does not explicitly find a link between Twitter bots and academic articles. We have taken several features from Altmetric, Twitter, and Botometer and fed them into a logistic regression model. We achieved 70% accuracy in predicting spam activity on any scholarly article on Twitter. We have also analyzed and concluded that there is significantly higher bot activity in the health and human science discipline than in other research areas.
CHAPTER 5

CONCLUSION

In this study, we have explored Twitter data to understand COVID-19-related topics and (de)motivation toward COVID-19 vaccination. We analyzed the temporal evolution of topics and the change in public stance. Here we discuss the summary and future works for our project.

5.1 Summary

In the first section of the thesis, as explained in Chapter 2, we analyzed (de)motivating topics and public stance. Although there are several works on identifying misinformation and topics related to COVID-19, our work focuses on identifying (de)motivating topics for COVID-19 vaccination. There is very little work in this segment, to our knowledge. We also identified the stance of the tweets and their geographic locations. Analyzing the tweet stance in contrast to their geographic location and primary topic conveyed intriguing results. We noticed that while demotivating topics differ based on geographic location, motivating topics remain mostly the same. This work can be crucial for future crises or public health emergencies to address public welfare and convey the message by cutting through the noise on social media.

In Chapter 3 we developed an interactive visualization tool to explore the data and understand the relation between the (de)motivation topics and public stance. We identified the prominent topics that are (de)motivating the public and noticed that some tweets,
even in favor of vaccination, can demotivate the public. We also noticed that politics and religion were more resonating topics in demotivating tweets, whereas factual statistics, education, and such were more prominent in motivating topics. This work has the potential for different research areas to interpret scientific data in an interactive method and bridge the gap between the public and the scientific community.

In Chapter 4, the last part of the thesis, we analyzed the Twitter data from Altmetric to understand the extent of spamming by Twitter bots on academic articles. We found that articles in the health science field are more prone to spamming and can be misrepresented on social media by bots to spread misinformation and distrust. Further exploration in this area can help identify and stop the use of scholarly articles to spread misinformation and cause distrust toward science. While distrust toward the scientific method is one of the significant reasons for demotivation toward COVID-19 vaccination [7, 8], misrepresenting scholarly articles to embolden the distrust is more harmful to the advancement of science and for motivating the public about health emergencies.

5.2 Future Work

Our work opens up different routes for future exploration. A more granular analysis of the topics can identify the topics important to different ethnicities, financial states, and other demographics of people. While we have created a stratified sample of the whole Twitter data source, analyzing the whole dataset can identify more niche topics and lead to better understanding.
We can investigate different visualization methods to improve usability, assisting in easier information consumption for individuals without scientific knowledge. This process can help rebuild trust in the scientific method as well as exploring the usefulness of visualizations and recommendation systems [169, 170, 171].

Finally, we can expand our work in identifying spamming in scientific articles to discover the research that Twitter bots are spamming to demotivate people about COVID-19 vaccination. A similar method can help any other field in the future.
REFERENCES


[20] Altmetric. What are altmetrics?
Accessed: 2021-1-5.


[29] Leysia Palen. Online social media in crisis events.
https://er.educause.edu/articles/2008/8/online-social-media-in-crisis-events,


[159] Jose Luis Ortega. The presence of academic journals on Twitter and its relationship with dissemination (tweets) and research impact (citations), 2017.


APPENDIX A

APPENDIX FOR CHAPTER 2
A.1 Datasets

A.1.1 Labeled Twitter Dataset

The dataset contains tweet IDs along with the location and tweet timestamp. The tweets are labeled based on motivating/demotivating status, stance towards COVID-19 vaccine, and topic in the tweet text. To comply with Twitter guidelines, we removed the tweet texts and author information. You can use Hydrator API [95] to hydrate the tweets. The anonymized dataset is available at:

https://github.com/DATALab-NIU/COVID-19-Twitter-Demotivating-Topics

A.2 Machine Learning Models

A.2.1 (De)Motivation Classifier

A pre-trained text classifier model to label the tweet texts as motivating or demotivating. The model is available at:

https://github.com/DATALab-NIU/COVID-19-Twitter-Demotivating-Topics

A.2.2 Vaccine Stance Classifier

A pre-trained text classifier model to identify the stance of a tweet text towards COVID-19 vaccine.
The model is available at:
https://github.com/DATALab-NIU/COVID-19-Twitter-Demotivating-Topics

A.2.3 Topic Model

A pre-trained topic model based on BERTopic to identify the topics in tweet texts. The models for demotivating and motivating topics are available at:
https://github.com/DATALab-NIU/COVID-19-Twitter-Demotivating-Topics
APPENDIX B

APPENDIX FOR CHAPTER 3
B.1 Visualization

B.1.1 Visualization Tool

The visualization tool is available at:
https://ashiqur-rony.github.io/visualize-covid-stance/

B.1.2 Source Code

The source code for the visualization is available at:
https://github.com/ashiqur-rony/visualize-covid-stance
APPENDIX C

APPENDIX FOR CHAPTER 4
C.1 Datasets

C.1.1 Labeled Twitter Dataset

The dataset contains Tweet IDs along with the bot score. To comply with Twitter guidelines, we removed the tweet texts and author information. You can use Hydrator API [95] to hydrate the tweets.

The dataset is available at:
https://github.com/DATALab-NIU/Twitter-Bot-Spamming-on-Academic-Article

C.2 Machine Learning Model

C.2.1 Spam Prediction Model

A pre-trained logistic regression model to predict the possibility for an article to be spammed or not.

The model is available at:
https://github.com/DATALab-NIU/Twitter-Bot-Spamming-on-Academic-Article