# Content Diffusion from Social Media to News Quotes

## Master Thesis

presented

by

**Raigon Kunnath Augustin**

at the

Universität Konstanz

Faculty of Sciences
Department of Computer and Information Science

1. Evaluated by Jun.Prof. Dr. Andreas Spitz
2. Evaluated by Dr. Johannes Fuchs

Konstanz, 2023

Raigon Kunnath Augustin

# Content Diffusion from Social Media to News Quotes

# Abstract

The process of news gatekeeping plays a pivotal role in influencing public opinion on various matters, encompassing not only the selection of information but also impacting the nature and substance of disseminated news messages. The emergence of social media platforms such as Twitter and Facebook has brought about a transformation in the journalistic landscape and the consumption patterns of the general public. These platforms have facilitated the easy sharing of opinions by both the general public and prominent figures in society, leading to a growing dependence of journalists on social media for content creation. Furthermore, these platforms have gained significant traction during recent social movements, exemplified by the widespread adoption during social movements like #MeToo. In this study, our objective is to investigate the diffusion of content from Twitter to online news quotes and examine the resultant shifts in the news gatekeeping process within this framework. To accomplish this, we collected tweets related to social movements from Twitter and employed Quotebank to identify tweets that were reported by news outlets as direct quotations. By leveraging diverse features from Twitter, Quotebank, and other data sources, as well as employing state-of-the-art natural language processing techniques, we analyze the changes in news gatekeeping. Our study revealed a low number of reported social movement-related tweets, and we observed that tweet characteristics indicating public engagement were associated with a higher likelihood of being reported. Furthermore, our findings suggest that journalists prioritize audience preferences when reporting tweets from celebrities, but they may consider their political bias when reporting tweets from politicians. This work offers valuable insights into social movement tweet coverage in news media, tweet selection factors, author selection bias, trends in reporting social movements, objectivity in reporting, and negative news coverage.

# Acknowledgment

I would like to express my sincere appreciation to my advisor, Jun. Prof. Dr. Andreas Spitz, for providing continuous feedback that proved invaluable throughout the conceptualization and writing process of this thesis. Your insights and ideas greatly assisted me in making crucial design decisions. I am grateful for your openness, patience, and the long discussions we had. Working with you on this project has been a tremendous learning experience, expanding my horizons in various ways. Additionally, I am thankful to Dr. Johannes Fuchs for accepting the role of the second reviewer for my work.

I received assistance from Tobias Nusser in the data collection process from Twitter, which greatly expedited the data collection timeline. I am immensely grateful for his valuable support.

I would like to extend my heartfelt gratitude to my family, who have been a constant source of support and encouraging me during the most challenging moments. I am deeply thankful for your unwavering belief in me.

# Table of Contents

# List of Figures

# List of Tables

# Abbreviations

**ANOVA** Analysis of Variance

**BERT** Bidirectional Encoder Representations from Transformers

**BFGS** Broyden–Fletcher–Goldfarb–Shanno

**BiLSTM** Bidirectional Long Short-Term Memory

**HSD** Honestly Significant Difference

**LCS** Least Common Substring

**LIWC** Linguistic Inquiry and Word Count

**LSTM** Long Short-Term Memory

**MLM** Masked Language Modeling

**MFOL** March For Our Lives

**NLP** Natural Language Processing

**NSP** Next Sentence Prediction

**RNN** Recurrent Neural Network

**SVM** Support Vector Machine

**TSC** Target-dependent Sentiment Classification

**QID** Query ID

# CHAPTER 1
# Introduction

The news media's coverage is limited to a small fraction of the multitude of events that take place globally every day. This set of selected news plays a significant role in shaping our perception and comprehension of the world [3]. The news media's focus on particular topics over others can impact our perception of the importance of these issues as members of the public. Additionally, even the decision-making of governing authorities and institutions may be swayed by the level of attention and coverage given to specific issues by the news media [19]. Hence News Gatekeeping is a crucial process in modern public life. [62] defines News Gatekeeping as "the process by which countless occurrences and ideas are reduced to the few messages we are offered in our news media". The gatekeeping process has a crucial role in shaping the public perception of different issues, as it not only involves selecting information but also affects the nature and content of the news messages that are disseminated. How we perceive our lives and the world around us is shaped by the gatekeeping process, which ultimately influences the social reality of each individual. In the past, these gatekeepers used to be editors, reporters, or other media personnel. And this selection process was not objective and involved several subjective factors such as the gatekeeper's own beliefs and assumptions, resource availability, and much more [62]. In their study regarding news production, [60] observed that gatekeepers in the news industry are frequently burdened by their work demands, and are influenced by the values and priorities of their employers, which are seen as an integral aspect of the news ecosystem. Frequently, the policy was established by the publishers and adhered to by the reporters [12]. However much has happened in the new media landscape in the last two decades due to the new technological advances. The emergence of social media platforms such as Twitter and Facebook has led to the creation and adoption of novel approaches, techniques, and resources for generating, disseminating, and sharing news content, both within and beyond the traditional newsroom [13].

## 1.1 Impact of Social Media on News Media Landscape

In recent decades, the landscape of news media has undergone significant transformations, transitioning from traditional formats such as newspapers, TV news programs, and radio stations to the realm of online news websites and social media platforms. The advent of advanced technologies and the widespread adoption of social media have further catalyzed substantial shifts in the way the public consumes news and these social media platforms have become the source of news for a large number of users [42]. News media organizations and journalists have embraced these evolving dynamics and adjusted their strategies accordingly. The majority of news organizations have recognized the significance of online and social media platforms and have established their digital presence [51, 41]. A survey conducted among journalists in 2017 revealed that nearly half of the respondents (48%) expressed their dependence on social media as an indispensable tool for their professional work [71]. Social media users play a crucial role in sharing online news articles, significantly contributing to the amplification of readership for news organizations [5]. If media organizations aim to expand their readership and if sharing news leads to increased readership, then the editorial choices of organizations will be influenced by the preferences of the audience [65]. As a result, journalists and news organizations are now more responsive to user preferences, as journalists and editors are incentivized based on positive digital metrics [8, 10, 73]. Additionally, social media platforms such as Twitter and Facebook have facilitated the effortless expression of public opinions on diverse subjects. Furthermore, individuals in positions of power and societal elites utilize these platforms to voice their opinions, contributing to the increasing reliance of journalist on social media for content creation as well [71].

In light of these recent advancements, there is a need for a comprehensive investigation to examine the utilization of social media platforms by journalists as content sources and to understand how audience preferences and social media have reshaped traditional news gatekeeping in the media landscape. Previous research [71] has shown that Twitter, in particular, is frequently employed as a source by news outlets. In this thesis, we will analyze the dissemination of content from Twitter to online news quotes and explore the changes to the news gatekeeping process in this context.

## 1.2 Twitter and Social Movements

The advent of social media platforms like Twitter has provided individuals from diverse backgrounds with a convenient means to voice and disseminate their opinions. Consequently, social movements have increasingly found a virtual stage on these platforms, shifting the locus of activism from physical streets to digital realms. The increasing adoption of digital social media platforms by participants in social movements represents a transformative trend in the communication landscape of social protests [64]. This phenomenon is widely recognized as a significant factor in the successful mobilization of recent social movements such as #MeToo and March For Our Lives(MFOL). The utilization of dedicated hashtags to share social movement-related information and the personal testimonies of movement participants have played a pivotal role in energizing these movements, elevating them to subjects of broader societal discourse [22].

Twitter platform played a pivotal role in facilitating the #MeToo movement, which emerged in response to the widespread revelations of sexual harassment in the film industry. The movement gained momentum in October 2017 when actress Alyssa Milano used the hashtag #MeToo to encourage individuals to share their experiences. The hashtag quickly became viral, resulting in over 4.5 million posts across various social media platforms within a day [33]. Over the course of a year, the hashtag was utilized an astonishing 19 million times [25].

In recent decades, the United States has experienced a series of highly publicized mass shootings [45]. Following a tragic school shooting on February 14, 2018, at Marjory Stoneman Douglas High School in Parkland, Florida, USA, where 17 students and teachers lost their lives, a group of student survivors swiftly responded by initiating an anti-gun violence movement known as March for Our Lives. Led by these students, the movement promptly advocated for legislative reforms concerning gun control. To amplify their message, student activists from Parkland, including Emma Gonzalez, David Hogg, and Cameron Kasky, leveraged social media platforms such as Facebook and Twitter using the hashtags #MarchforourLives and #NeverAgain to raise awareness about the issue of gun control and orchestrate a nationwide rally [23]. The MFOL rally in Washington D.C drew an impressive crowd of approximately half a million people, while simultaneously over 800 rallies were coordinated in various locations across the United States [80].

As discussed before, media emphasis on certain issues and not other issues determine which issues we as members of the public think are important [19]. Hence for social movements like #MeToo and #MFOL to be successful in terms of public awareness, and action against accusers, the public has to be made aware of such movements in the right context. More often how the public perceives social movements like MeToo is largely dependent on how these movements are covered by the news media [3]. Even the decision-making of governing authorities and institutions can be influenced when such movements are given more attention and coverage by the news media [19]. Building upon the extensive utilization of Twitter observed during the #MeToo and #MFOL movements, this study focuses on examining the diffusion of #MeToo and #MFOL-related tweets from Twitter to news quotes in greater detail. To measure this content diffusion, we are planning to use Quotebank [70] dataset which is a web-scale corpus of over 235 million unique speaker-attributed quotations extracted from 196 million English news articles.

Based on previous research highlighting the news gatekeeping process, it has been noted that events involving prominent figures and celebrities often receive considerable attention [31]. Therefore, in our study, we specifically concentrate on analyzing social movement-related tweets shared by Hollywood celebrities and U.S. politicians, hereafter referred to as the **Celebrities** and **Politicians** user groups, respectively.

## 1.3 Research Objectives and Contributions

This thesis aims to examine the dissemination of social movement-related tweets by Celebrities and Politicians to news quotes. Specifically within the context of social movements, our objective is to assess the extent to which journalists utilize social media platforms like Twitter as a source for their news content. Additionally, we seek to explore the influence of factors such as audience preference and author popularity have on the news gatekeeping process. Furthermore, we investigate whether traditional news media behaviors are reflected in the reporting of tweets as direct quotations. Given the potential presence of political biases in news media outlets, we also conduct various analyses to gain insights into the reporting behavior of these outlets regarding social movements.

Below are the research questions explored in this thesis:

- **Research Question 1:** What is the relative influence of tweet-level and author-level attributes on the gatekeeping process for the selection of tweets related to social movements in the news media, and which of these attributes has the greatest impact on the likelihood of a tweet being published?
- **Research Question 2:** To what extent do news outlets exhibit a preference for reporting sub-tweets that convey negative emotions expressed by the author?
- **Research Question 3:** To what extent do news outlets exhibit bias in their selection of authors to report on tweets in the context of social movements?
- **Research Question 4:** How did the political bias of news outlets influence the reporting of social movement- related tweets?
- **Research Question 5:** How do news outlets portray the authors of social movement-related tweets they publish?

The git repository [1] contains the source code necessary for reproducing the experiments and results presented in this thesis.

## 1.4 Structure of the Thesis

This thesis introduces the impact of social media on the news gatekeeping process and focuses on the need for meaningful investigations to analyze the diffusion of content from social media to new quotes and explore the changes to the news gatekeeping process in this context. Therefore, in Chapter 2 we provide a detailed overview of the literature about news gatekeeping and how social media is changing the news media landscape. Additionally, in Chapter 2 we provide background

[1] `https://gitlab.inf.uni-konstanz.de/raigon.kunnath-augustin/content-diffusion`

information regarding various techniques that will be used during our experiments. Chapter 3 contains a discussion about data sources that have been utilized for our experiments and also outlines the data processing steps involved in the creation of the dataset. The research questions and hypothesis that will be explored in this thesis, and the methodology used to implement our experiments are discussed in Chapter 4. Additionally, Chapter 4 contains implementation details and results of our experiments. In Chapter 5, we provide a detailed discussion of these findings, along with an exploration of the limitations inherent in our current study. With Chapter 6, we conclude the thesis by providing a summary of our contributions to understanding the content diffusion from social media to news media.

# CHAPTER 2
# Background and Related Work

## 2.1 Background

This section offers an overview of the technical concepts employed in the implementation of this thesis. We present the Natural Language Processing (NLP) methods employed in this thesis in subsection 2.1.1. Additionally, in subsection 2.1.2, we discuss the text similarity measures used to measure content diffusion from Twitter. The statistical analysis techniques implemented for our experiments are explained in section 2.1.3, while section 2.1.4 provides a comprehensive overview of the machine learning models utilized for classification tasks.

### 2.1.1 NLP Methods

In this subsection, we present the NLP methods and models used in this thesis.

#### 2.1.1.1 Transformer-based Language Models

Transformer [69] architecture introduced in 2017 sparked significant innovations in the field of NLP. It catalyzed advancements and breakthroughs in various NLP tasks. Transformers rely on self-attention mechanism, allowing the model to capture the dependencies between words in an input sentence more effectively. Self-attention mechanism to a large extent solved the long-range dependencies problems of RNNs [58, 75] and LSTMs [35]. Additionally, transformers use positional encoding to incorporate the sequential information of words into the model. The attention mechanism enables the transformers to assign varying levels of importance to different words in an input sequence, enabling it to focus on the most relevant information for a given task.

**BERT**

Bidirectional Encoder Representations from Transformers (BERT) [21] is a transformer-based language model introduced by Google in 2019. It has made a major impact on the field of NLP by changing how we approach language understanding and generation tasks. BERT was pre-trained on two tasks Next Sentence Prediction (NSP) [21] and Masked Language Modeling (MLM) [21] which have allowed it to learn contextualized word representations. By training on NSP and MLM tasks, BERT was able to generate contextualized word representations that capture various aspects of language semantics and syntax. BERT's architecture comprises multiple transformer layers, facilitating the capture of local and global dependencies within the input text. Operating bidirectionally, BERT effectively utilizes contextual information from both the left and right contexts of a word, enabling a comprehensive understanding of its semantics. Upon its introduction, BERT demonstrated exceptional performance, surpassing previous benchmarks and establishing itself as a state-of-the-art model across various NLP tasks.

In this thesis, we have utilized several variants of BERT for conducting various experiments. Further elaboration on these models will be provided in the subsequent sections of this chapter.

#### 2.1.1.2 Perspective API

Perspective API[1], developed by Google, is a machine-learning model specifically trained to assess the potential toxicity of textual content. By leveraging multilingual BERT-based models and training on a vast amount of online comments from diverse platforms like Wikipedia, Perspective API can generate a score between 0 and 1 that represents the likelihood of the text being perceived as offensive, disrespectful, or harmful.

The API offers this score for the following six attributes: **(1)** Toxicity, **(2)** Severe Toxicity, **(3)** Identity Attack, **(4)** Insult, **(5)** Profanity and **(6)** Threat.

#### 2.1.1.3 TweetNLP

TweetNLP [15] is an integrated python library [2] for social media analysis. It supports a diverse set of NLP tasks such as emotion detection, sentiment analysis, and much more. All the task-specific models are powered by RoBERTa [44] language model fine-tuned on Twitter data. In this thesis we use the TweetNLP package for the following four NLP tasks:

- **Sentiment Analysis:** The objective of the sentiment analysis task is to classify tweets into one of three categories: positive, negative, or neutral. Semeval2017 dataset for Subtask A [57] was used to fine-tune the model for this task.

[1] `https://perspectiveapi.com/`
[2] `https://tweetnlp.org/`

- **Emotion Detection:** The objective of this task is to identify the emotion elicited by a tweet. The model was further fine-tuned on the emotion detection task by using the SemEval2018 "Affects in Tweets" [46] dataset.

- **Hate Speech Detection:** The objective of this task is to predict whether a tweet exhibits hate speech towards two specific target communities: immigrants and women. SemEval2019 Hateval challenge dataset [7] was used to fine-tune the model.

- **Offensive Language Detection:** The objective of this task is to identify the presence of offensive language in a tweet. SemEval2019 OffensEval dataset [79] was used for model fine-tuning.

Both the TweetNLP library and Perspective API will be used in section 3.3 of Chapter 3 to generate tweet-level features.

#### 2.1.1.4 Author Emotion Detection

In the field of NLP, author emotion detection involves the automatic identification and classification of the emotional state expressed by an author in a given text. The recent advancements in this task have been facilitated by the effectiveness of pre-trained language models like BERT [21] and the availability of extensive and well-annotated datasets [2]. In contrast to sentiment analysis, emotion detection focuses on identifying specific emotions expressed by the author and can often involve a large set of target classes. Below we discuss the various emotion detection models we have used in this thesis.

- **Model 1:** This emotion detection model [2] utilizes a neural network architecture combining BERT [21] and Bi-LSTMs [26]. The model is fine-tuned on the emotion detection task using the GoEmotions dataset [20]. Figure 1 illustrates the architecture of the model. The GoEmotions dataset consists of 58,009 English text snippets collected from Reddit, encompassing 28 different target emotion classes. In this model, the input text is inputted into BERT, and the resulting word embeddings from the last layer are passed through the Bi-LSTM network. The output from the Bi-LSTM is then pooled together to obtain the target label.

- **Model 2:** This emotion detection model [54] employs an adapter-based training strategy that achieves comparable performance to full model fine-tuning, while significantly reducing the number of parameters required [36]. Adapters are comprised of a compact set of freshly initialized weights incorporated at each layer of the transformer. During fine-tuning, these weights are trained while the pre-trained parameters of the large language model remain fixed [53]. For model 2[3], an adapter for the BERT-BASE [21] model was trained using

[3] `https://adapterhub.ml/adapters/AdapterHub/bert-base-uncased-pf-emo/`

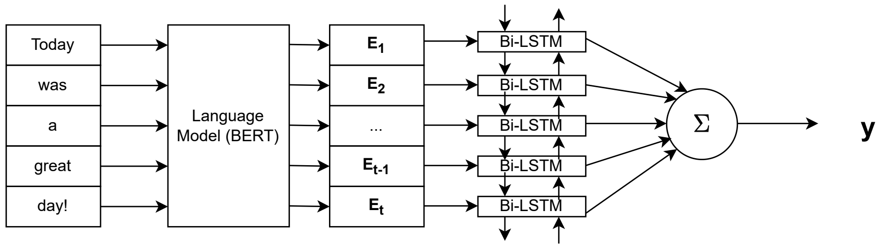


**Figure 1** Bi-LSTM Architecture used by model 1 [2] for emotion detection. Here $\mathbf{E}_n$ are word embeddings for the input text from BERT and $\mathbf{y}$ is the predicted emotion label

the EmoContext dataset [17] and incorporates a classification head for predicting the emotion. The EmoContext dataset, widely adopted as a benchmark in emotion detection, comprises text conversations annotated with emotion labels. Each conversation includes a contextual snippet and an associated response, providing valuable data for training and evaluating emotion detection models. It has 4 emotion labels: joy, sadness, anger, and others.

- **Model 3:** Model 3 [32] employs a DistilRoBERTa-base model which was fine-tuned using 6 distinct emotion detection datasets encompassing a wide range of text genres. This model encompasses 7 emotion labels, namely joy, sadness, anger, disgust, fear, neutral, and surprise.
- **Model 4:** Model 4 is not a machine learning model. It uses Linguistic Inquiry and Word Count (LIWC) [52] for emotion detection. LIWC is a text analysis tool that employs a pre-existing dictionary of words categorized into different linguistic and psychological groups. This tool can also be employed to ascertain the emotional content of a given text. To determine the emotion of a text using LIWC, we calculate the total count of positive emotion words and negative emotion words within the text. The difference between these two counts is then computed. A positive result indicates the presence of positive emotion in the text, and a negative result suggests the presence of negative emotion. If the result is zero, it signifies a neutral emotional tone in the text.

The experiments outlined in the section 4.2 of Chapter 4 will utilize all the emotion detection models introduced above.

#### 2.1.1.5 Target-dependent Sentiment Classification (TSC)

TSC aims to classify the sentiment or opinion expressed towards a specific target entity within a given text. It involves identifying the sentiment as positive, negative, or neutral considering the context of the target entity mentioned in the text. The depiction of individuals in news coverage of political topics holds great importance, as it significantly influences both individual and societal opinion formation processes

[9]. Hence NLP tasks such as TSC and stance detection play a crucial role in understanding perspectives expressed in text data.

In this thesis, we are using the TSC model introduced in [27]. Their model has 4 key components: a pre-trained language model, a representation of external knowledge source (EKS), a target mention mask, and a bidirectional GRU [18]. As the pre-trained language model they have used RoBERTa model [44] and as EKS they have used SENT [37] (a sentiment dictionary), MPQA [76] (a subjectivity dictionary) and NRC [47] (a dictionary of sentiment and emotion). By combining the contextualized word embedding from language models and the knowledge representation from EKS, they were able to get good performance for the TSC task. Their model was fine-tuned on the NewsMTSC dataset [27] which is a manually annotated dataset for TSC in political news articles. Their training task is very similar to our *research question 4* which investigates how news media portrays politicians in news articles. The NewsMTSC dataset has three target class labels: Positive, Negative, and Neutral.

Below we discuss the preprocessing steps we have used for using the above-discussed TSC model:

- **Step 1:** Get the speaker name from Quotebank. [70].
- **Step 2:** Combine the left and right context in Quotebank (more details in Chapter 3) to create the news article snippet.
- **Step 3:** Replace the first and last names in the article snippet with the full name of the author.
- **Step 4:** Perform coreference resolution[4] in the article snippet for speaker name.
- **Step 5:** Find the first occurrence of the target (speaker name) in the article snippet and split the snippet: Left, target-mention, Right.

The experiments outlined in the section 4.4 of Chapter 4 will utilize the TSC model and the preprocessing steps discussed above.

### 2.1.2 Text Similarity Measures

Quantifying the semantic and lexical similarity between multiple texts is a pivotal task where similarity measures assume a vital role. A plethora of text similarity measures is at our disposal, employing diverse techniques. The selection of an appropriate text similarity measure depends on the specific NLP task and the inherent characteristics of the data. Choosing the right method is crucial to ensure accurate and meaningful results are obtained. In this thesis, we are using the text similarity

[4] https://github.com/huggingface/neuralcoref/releases/tag/v4.0.0

method to measure the content diffusion from tweets to news quotes. For this, we need to accurately measure the amount of overlap between tweet text content and quote text content. In the below sub-sections, we briefly discuss the text similarity measures we have explored to achieve this.

#### 2.1.2.1 Jaccard Similarity

Jaccard is a similarity measure used widely in machine learning and NLP to quantify the similarity between two sets of elements [43]. It is useful for comparing the similarity of text segments based on their shared tokens. It quantifies the ratio of shared elements between the two sets in relation to the overall count of unique elements. The similarity score ranges between 0 and 1, with 1 indicating identical sets and 0 indicating no similarity between the sets.

Jaccard similarity between two texts can be computed as,

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

In the above formula for computing Jaccard similarity, $A$ and $B$ represent the sets of unique words or tokens in the two texts being compared. The intersection of A and B is denoted by $A \cap B$, and the union of A and B is denoted by $A \cup B$. The vertical bars surrounding the sets represent the cardinality of each set, indicating the number of elements in each set.

#### 2.1.2.2 Cosine Similarity

In NLP cosine similarity is widely used to measure the similarity between two text segments. It measures the similarity based on the angle between the two vectors which represents the text segments. It is useful for comparing text documents with varying lengths as it utilizes the direction rather than the magnitude of the vectors. In this thesis, we can convert the tweet text content and the news quote text content to sentence embeddings using BERT [21] or any other machine learning models and use cosine similarity to measure the similarity between them. The cosine similarity score ranges between 0 and 1, where higher values indicate greater similarity between the text segments.

Cosine similarity between two texts can be computed as,

$$\text{Cosine Similarity}(A, B) = \frac{A \cdot B}{\|A\|\|B\|}$$

In the above formula, $A$ and $B$ represent the two word embeddings being compared. $A \cdot B$ denotes the dot product between the two embeddings. $\|A\|$ and $\|B\|$ represent the Euclidean norms embeddings.

#### 2.1.2.3 Lest Common Substring

LCS (Least Common Substring) is another commonly employed approach for assessing the similarity between two segments. It identifies the shortest substring that is shared between the two text segments. By determining the length of the LCS, we can gauge the degree of similarity between the segments. A greater LCS length in relation to the original length of the text segments indicates a higher level of similarity. This method proves valuable in NLP tasks that necessitate quantifying the extent of overlap between text segments.

$$\text{LCS Similarity}(A, B) = \frac{\text{length of the longest common substring of } A \text{ and } B}{\text{length of } A + \text{length of } B}$$

In this formula, $A$ and $B$ represents the two strings to be compared.

#### 2.1.2.4 Levenshtein Distance

Levenshtein distance, also known as edit distance, is a widely used text similarity measure in the field of NLP. It takes into account both lexical and structural differences between text segments. The similarity is determined by counting the minimum number of operations required to transform one text segment into another. These operations include character insertions, deletions, and substitutions. When calculating the Levenshtein distance, each character in one segment is compared with the corresponding character in the other segment. If the characters do not match, we can choose to insert, delete, or substitute the character to achieve a match. The similarity is then computed as the sum of the operations necessary to transform one segment into the other.

### 2.1.3 Statistical Analysis for Observational Study

The abundance of big data in today's era signifies a transition from a time when behavioral data was scarce to a time when it is abundant, owing to the constant influx of data. A lot of the big data created can be considered observational data. According to [59] observational data refers to data that is obtained by observing a social system without any form of intervention, in a general sense. The main objective of an observational study is to gain insights into real-world phenomena, understand the relationships between variables, and make observations about how certain factors may be associated with specific outcomes. All the experiments conducted in this thesis can be categorized as observational studies since we are working on observational data from various data sources. The statistical methods discussed below are used in the experiments discussed in Chapter 4.

#### 2.1.3.1 Logistic Regression Analysis

The primary aim of conducting logistic regression analysis is to develop a model that predicts the probability of a binary outcome by considering the values of the independent variables. This enables us to gain insights into the influence of different predictors on the likelihood of an event taking place. Moreover, logistic regression allows us to estimate the coefficients associated with each independent variable, providing information about the direction and magnitude of their impact on the probability of the outcome. One notable advantage of utilizing logistic regression in an observational study is its capability to control for confounding variables by incorporating them as independent variables in the model.

#### 2.1.3.2 Propensity Score Matching

In randomized controlled experiments, treatment or control assignment is determined by a random process such as flipping a coin or dice, while in observational studies, treatment assignment is not controlled experimentally [56]. In randomized experiments, this randomization process effectively balances the observed covariates, to some extent, and offers a foundation for anticipating that unmeasured covariates will also tend to exhibit a similar balance [56]. Since randomization is not used to assign treatments, confounding variables may potentially impact the desired outcome in an observational study. To mitigate this concern, propensity score matching strives to establish comparable groups with similar distributions of the observed covariates. By reducing the imbalance in observed covariates between the treatment and control groups, propensity matching enhances the reliability of comparisons and yields a more accurate estimate of the causal effect of the treatment.

In this thesis, we have calculated the propensity score by using a logistic regression model, where the treatment assignment is the dependent variable and the observed covariates are the independent variables. The predicted probability by the logistic regression model is used as the propensity score. We identify instances in the control group that have a similar propensity score to each instance in the treatment group. To perform this matching we have used the nearest neighbor matching with 25% of standard deviation of the propensity score as the radius for matching.

#### 2.1.3.3 Wilcoxon Signed Rank Test

Wilcoxon signed-rank test [77] is a non-parametric statistical test used for the analysis of matched pair data. It can be utilized to compare the distribution of the paired differences in covariates after matching. This makes it a valuable tool for performing post-matching analysis after propensity score matching to determine the statistically significant predictors for the treatment.

#### 2.1.3.4 One-way ANOVA Test

The one-way ANOVA (Analysis of Variance) test is a statistical method used to compare the means of three or more groups to determine if there are any significant differences among them. In this test, the null hypothesis assumes that there are no significant differences among the groups. And the alternative hypothesis suggests that at least one of the groups differs from the others. In the context of time series data, a one-way ANOVA test can be used to assess whether there are statistically significant differences in the means of multiple time series. A statistically significant result indicates that at least one time series differs significantly from the others. After obtaining statistically significant results from the one-way ANOVA test, we can proceed with posthoc analyses, such as the Tukey's Honestly Significant Difference (HSD) test [1], to identify the specific time series that contribute to the observed differences.

#### 2.1.3.5 Tukey HSD Test

Tukey's Honestly Significant Difference (HSD) test [1] is commonly employed for comparing multiple groups. It is utilized following an ANOVA test to identify significant differences among the means of the groups. This test is employed when the null hypothesis of the ANOVA test is rejected. It conducts pairwise comparisons between all possible combinations of groups to identify specific pairs that exhibit significant differences.

### 2.1.4 Classification models

In this thesis, several machine learning models were trained for binary classification tasks in certain experiments. Considering the limited availability of data, the decision was made to exclude deep learning models from the analysis. The subsequent subsections provide a comprehensive discussion of these employed models. These models will be used in the experiment defined in section 4.1 of Chapter 4.

#### 2.1.4.1 RandomForest

The Random Forest algorithm is predominantly employed for classification tasks in machine learning. It operates as an ensemble method by combining numerous decision trees, each trained on a random subset of the training data. The final prediction of the Random Forest model is obtained by aggregating the predictions made by each tree. Random Forest demonstrates a reduced susceptibility to overfitting compared to standalone decision trees. Moreover, it offers built-in techniques for estimating the most influential features involved in the classification process.

#### 2.1.4.2 Support Vector Machines

Support Vector Machines (SVM) is a versatile machine learning algorithm suitable for classification tasks, particularly adept at handling high-dimensional data. The primary objective of SVM is to identify a hyperplane that maximizes the margin between two classes, facilitating optimal separation. To address non-linear data, SVM employs the kernel trick, enabling the transformation of input training data into a higher-dimensional space (kernel space) where linear separability is achieved.

#### 2.1.4.3 Gradient Boost

Gradient Boosting is a robust machine learning technique extensively employed in classification tasks. It utilizes an ensemble approach by combining multiple weak classifiers to form a potent predictive model. Specifically designed for classification, Gradient Boosting sequentially constructs an ensemble of decision trees, with each subsequent tree aimed at rectifying the errors made by its predecessors. Through this iterative process, it effectively reduces errors and enhances the overall accuracy of predictions.

#### 2.1.4.4 AdaBoost

AdaBoost, which stands for Adaptive Boosting, is a widely utilized machine learning algorithm for classification tasks. Like Gradient Boosting, it is an ensemble method that combines multiple weak classifiers to form a powerful predictive model. Each weak classifier is specifically trained to rectify the misclassifications made by preceding classifiers. Throughout the iterative process, AdaBoost assigns greater weights to the misclassified instances, enhancing their influence in subsequent iterations.

## 2.2 Related Work

In this section, we give an overview of the past works related to this thesis. The topic of this thesis can be segmented into two fundamental themes: content diffusion from Twitter to News quotes and changes caused by social media to the news gatekeeping process. In the next subsections, we will discuss the related work for each of these themes.

### 2.2.1 Twitter as News Source

The relevance of Twitter as a journalistic resource is growing, primarily due to its increased adoption by politicians and other influential individuals, as noted in a study by [50]. This can be attributed to the opportunity social media platforms provide for politicians to bypass the traditional gatekeeping role of legacy news media, allowing direct engagement with citizens, as highlighted by [24]. Journalists utilize Twitter in diverse ways, such as uncovering new stories, sourcing information and contacts, gathering quotes, and utilizing collective community knowledge for information verification, as mentioned in a study by [11]. In this thesis, our primary focus revolves around the utilization of Twitter by journalists as a source for acquiring quotes.

The previous research [50] that closely aligns with this thesis conducted a manual content analysis on news articles published in June 2019 within the political sections of three German newspapers and one German news magazine. They examined a total of 2,911 news articles and discovered that 8% of the overall news coverage included embedded tweets as direct quotations. Furthermore, they found that 5% of the total news coverage consisted of articles that paraphrased original tweets. The study also revealed that journalists predominantly used Twitter to cite authoritative sources and rarely utilized it to report tweets from the general public. To investigate the tweet characteristics that increase their likelihood of being reported, the researchers employed a Poisson regression model. They considered popularity cues such as the number of likes and retweets, as well as news factors like proximity, as independent variables. The dependent variable measured the frequency of individual tweets appearing across multiple news media outlets. The results indicated that popularity cues and proximity significantly predicted the inclusion of full tweets in political news coverage. However, it's important to note that their analysis did not take into account tweets posted by the authors that were not reported by any news outlets. In a similar vein, [14] conducted a manual content analysis on a five-year span of eight British and Dutch newspapers. Their analysis revealed the presence of 5,813 tweet quotes used as news sources in a total of 3,361 news articles. Notably, 92% of the tweets were reported as direct quotations, while the remaining tweets were paraphrased. The study also identified the top four sources utilized by journalists, which included celebrities, athletes, politicians, and the general public, accounting for approximately 79% of all the tweets cited. In another previous study [11] that examined the dissemination of content from tweets to newspapers, researchers conducted a manual content analysis focusing on the political news coverage section of Dutch newspapers. Their analysis revealed that journalists primarily cited tweets shared by politicians, accounting for 72% of all reported tweets. Furthermore, they discovered that 71% of the reported tweets were quoted verbatim.

In their study [49], the researchers examined the presence of reported tweets in online news articles sourced from German sports news providers. Out of a total of 3,150 online articles analyzed, they discovered that 507 articles (16%) relied on social media as a source, with Twitter being the predominant platform. Furthermore, through an analysis of article placement, they observed that news articles citing social media sources were more frequently positioned at the top of the page. [6] conducted a comparative analysis on the diffusion of content from Twitter to online news media and traditional print media. Their study involved analyzing a sample of randomly selected articles over three months. The findings revealed that online news media utilized Twitter as a news source to a higher extent than traditional news outlets, indicating a greater reliance on Twitter for news stories. In a similar study, [48] compared the utilization of Twitter as a news source among newspapers, TV networks, and cable channels. Their research revealed that TV networks relied on Twitter as a source more frequently than newspapers did. Furthermore, they observed that both print and broadcasting media predominantly used Twitter for reporting soft news rather than hard news topics. In their analysis, they did not find any significant correlation between the frequency of citing a tweet author and the number of Twitter followers that the author had.

Except for [49] and [6], the majority of the previous studies we discussed focused on utilizing print media as their data source for examining the diffusion of content from Twitter. Moreover, these studies had a limited scope in terms of the number of news outlets considered and did not account for the political bias of the outlets. According to [6], online news media websites exhibit a stronger inclination to utilize Twitter as a news source compared to print media. However, a comprehensive study that encompasses a wide range of news outlets with varying political biases is still lacking.

### 2.2.2 News Gatekeeping in Digital Media Landscape

As discussed in Chapter 1, the news gatekeeping process not only involves the selection of information but also influences the content and nature of the messages that are disseminated as news. With the introduction of new communication technology, traditional news gatekeeping has undergone changes. Past work has studied this at multiple levels.

According to [29] and [66] the combination of heightened exposure to audience feedback and a less stable financial landscape appears to be motivating journalists to adopt a more consumer-centric approach. This shift involves producing news that caters to people's preferences and interests, deviating from the traditional citizen-oriented perspective that emphasizes journalism's role in delivering information that people ought to know. Journalists can get audience preference and feedback through mainly three channels: Reader Comments, Social Media, and Web analytics [30]. In their research [30] examined the impact of exposure to and utilization of audience feedback mechanisms on the culture of journalism. Using data gathered from an online survey of 358 news journalists in Australia, their study discovered a positive correlation between the frequent reading of readers' comments and an amplified perception of the significance of both consumer-oriented and citizen-oriented approaches. Furthermore, they observed a connection between the perceived effectiveness of web analytics as a form of audience feedback and enhanced recognition of the importance of consumer orientation. Through 150 hours of observations and interviews with 31 journalists, the research work done by [67] found that the main use of social media in newsrooms can be grouped into three categories: monitoring, interacting, and promoting. They found both the editors and journalists monitor social media for topics that are trending and even consider them in planning stories for the website and even the physical paper. Adapting content to suit social media platforms exemplifies how audience influence manifests at the routine level. This occurs when journalists integrate audience behavior and preferences into their news production process.

Audiences now play a crucial role in the distribution of content within the online news media realm [67]. Most online news websites have a "share" feature that allows audiences to share particular articles they like with their friends and acquaintances through social media. In a related study, [5] found that social media users have a pivotal role in the dissemination of online news articles, playing a significant part in amplifying readership for news organizations. Additionally [65] observed that for increasing their readership, news organizations tend to give more importance to audience preference. In a related work [74] studied the impact of audience clicks in the news news selection process. Using a mixed-method approach, they examined both the print and online versions of five Dutch national newspapers. Their findings revealed that storylines featured in the most viewed articles were more likely to be covered in subsequent reports, suggesting that news selection is influenced by audience clicks.

Previous studies focused on examining the impact of social media on news gatekeeping have predominantly relied on surveys as their primary research methodology. According to [67], surveys targeting journalists offer valuable insights into the extent of social media usage within newsrooms. However, they fall short of providing a nuanced understanding of how social media are specifically employed in conceptually distinct ways that are relevant to the gatekeeping process. This limitation is partly due to the survey method. Apart from [50], none of the other studies attempted to explore the significance of tweet characteristics in determining whether a tweet gets reported. Even in the work by [50], only a limited number of factors such as popularity cues were considered while neglecting content-related features that could be extracted from the text of the tweets. Some of the limitations observed in previous studies can be attributed to their reliance on manual content analysis as the preferred method, without incorporating any NLP techniques.

## 2.3 Positioning of the Thesis

In this thesis, we investigate the dissemination of content from Twitter to news quotations in online news websites within the context of social movements. Our analysis encompasses a wide range of news outlets, taking into account their respective political biases which was missing from the previous research works. In this work we explore the multitude of tweet and author-level features offered by Twitter and also incorporate features from alternative data sources, aiming to gain insights into the news selection process. Rather than relying on surveys and manual content analysis, we employ cutting-edge NLP techniques and data-driven methodologies to comprehensively understand the dynamics of news gatekeeping. Moreover, we perform content analysis using NLP methods to investigate whether traditional news media behaviors are reflected while reporting tweets. To the best of our knowledge, this is the first study to examine the media coverage of social movements specifically in relation to the reporting of social movement tweets.

# CHAPTER 3
# Data

The first step in understanding the content diffusion from social media to news quotes is extracting and processing data from the right sources. In this chapter, we provide an overview of the various data sources we will be using, as well as the methods used for data extraction and the features of the data that are relevant to the experiments. Figure 2 presents an overview of the data extraction and processing procedures carried out in this study.

## 3.1 Data Sources

### 3.1.1 Wikidata

Wikidata [72] is an open knowledge base that contains data about various topics, peoples, places, and more. Each Wikidata item has a unique identifier called QID. For instance, people who have an entry in the Wikidata knowledge base will have a unique QID associated with them. Various useful data associated with that person such as age, nationality, gender, and more are stored under their Wikidata profile. As mentioned in Chapter 1, in our study we are focusing on tweets posted by U.S Politicians and Hollywood Celebrities, and considering the nature of their profession, we assume the majority of the users in these two user groups will have an entry in Wikidata.

Our methodology to retrieve the list of users from Wikidata is similar to [40]. For both the user groups we use a list of related occupations as a filter. The list of related occupations for the Politicians user group was obtained from [40]. The list of related occupations for the Celebrity user group, such as actors, and directors, was manually created for this study. Given that our analysis focuses on content diffusion from Twitter, we limited our user sample to those with a Twitter handle maintained in Wikidata. Accordingly, our study excludes users without a Wikidata entry or a Twitter handle listed in their Wikidata profile.

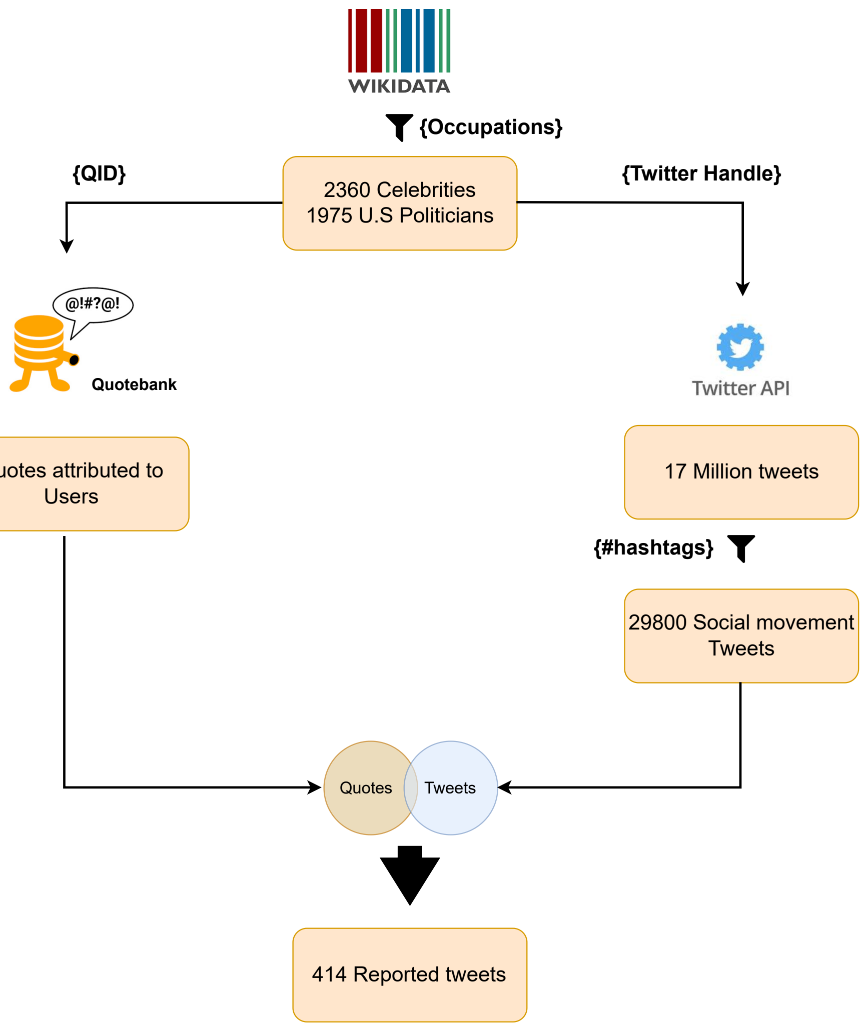


**Figure 2** An overview of the data extraction pipeline used to extract data from various data sources to create the dataset containing reported tweets.

### 3.1.2 Twitter

Twitter and other popular social media platforms allow people from all walks of life to express and disseminate their opinions regarding various topics. Consequently, social movements are now carried out on social media rather than on the streets. Twitter was extensively used during both the #MeToo and MFOL social movements. For this study, we have selected Twitter as the primary data source to extract the content generated by politicians and celebrities. In addition to the text content of tweets, Twitter offers various tweet-level and user-level attributes that aid in understanding why certain tweets are selected for news coverage. An example of a tweet-level feature that Twitter provides is the count of retweets, which represents the level of public engagement with that particular tweet. Later in this chapter, we will provide a detailed discussion of all such attributes.

On Twitter, hashtags allow users to efficiently distribute information and express viewpoints on social issues to a wide audience. Hashtag activism is a term coined to describe the use of Twitter hashtags for internet activism. Every social movement that leverages social media is associated with multiple hashtags. For instance, during the #MeToo movement, the hashtag #MeToo was used 19 million times in a single year [25]. Therefore, to retrieve the social movement-related tweets posted by users in our two user groups, we are utilizing a compilation of hashtags linked to the social movement. Some of the commonly used hashtags associated with the two social movements considered for our study are #MeToo, #TimesUp, #NoMoore, #marchforourlives, #guncontrolnow, #parkland, etc. We created this list of hashtags manually by referencing past similar works [25].

We are utilizing the Twitter Academic API to extract the tweets posted by the users. To prepare the tweet content for further analysis, we apply the following pre-processing steps to remove any unwanted content that may be present.

**Pre-processing steps:**

- **Remove hyperlinks:** URLs, consisting of arbitrary characters and symbols, are removed as they do not contribute to the overall meaning of the content.
- **Remove Emojis:** Emojis serve as graphical symbols employed to express emotions or depict visual elements within text. Although they enhance the visual aspect of communication, in certain NLP tasks primarily centered around textual analysis, they may be perceived as extraneous noise. Hence we remove them from the tweet text.
- **Removing tweets with less than 5 words:** Tweets with very few words might lack sufficient meaningful information and are hence not suitable to be used for further analysis. Most of the NLP-based techniques we have used in this thesis will perform poorly if tweets with very few words are used as input.
- **Converting to lowercase:** Lowercasing the text during preprocessing aids in normalizing the data, ensuring a consistent representation. By eliminating capitalization variations, it fosters a more standardized and uniform dataset, facilitating effortless comparison, matching, and analysis of the text.

### 3.1.3 Quotebank

Quotebank [70] dataset is a web-scale corpus of over 235 million unique speaker-attributed quotations extracted from 196 million English news articles. The news articles used for creating this dataset were provided by Spinn3r.com and these articles were published between September 2008 and April 2020. For creating Quotebank, a BERT [21] language model was fine-tuned to predict the speaker of a quotation in the news article from the text surrounding the quote. Speakers in Quotebank are identified by their unique QID from the Wikidata knowledge base.

There are two versions of the Quotebank dataset available: Quotation-centric and Article-centric version. Both these versions are publically available to download [1]. For this study, we will be using the article-centric version of Quotebank as it contains all individual quotation mentions with the associated speaker and the article context in which they are mentioned. The article-centric version offers multiple data fields which are relevant to our study. It has a **date** field which indicates the publication date of the article. The **quotation** field contains the text content of the quote reported in the news article. The field **leftContext** contains the text content that appears at the left context window of the quotation. Similarly **rightContext** contains the text content that appears at the right context window of the quotation. The **globalTopSpeaker** field contains the most probable speaker of the quotation. Additionally, the **URL** field contains the link to the original news article. We can extract the name of the news outlet which published the quotation from the URL field.

Tweets that have been reported in the news articles published between 2008 and 2020 should have a higher probability to be present in Quotebank as a speaker-attributed quotation. As discussed above, Quotebank provides the context in which a quotation was reported by the news outlet which helped us to understand how the tweet was reported by various outlets.

**Table 1** Five primary media bias categories from medibiasfactcheck.com and some of the news outlets in these categories

| Media Bias | News Outlets |
|---|---|
| Left | Arizona Mirror, Chicago Reader, LA Weekly, The Baffler |
| Left-Center | CBS News, Arizona Daily Sun, Insider, NBC News |
| Center | Capitol Fax, CNET, Newark Post, C-SPan |
| Right-Center | CCN, First Post, Times News, Bild |
| Right | The Sun, The Patriot Post, New York Sun, New Boston Post |

### 3.1.4 Media Bias / Fact Check

Media Bias/Fact Check [2] is a website that operates independently and offers a bias rating for various media sources. Their website currently lists over 6100 media sources and journalists. Media sources are classified into nine different bias categories by the website, each representing a different political leaning or perspective. Following the approach of previous studies [38], we limit our analysis to the five primary bias categories. The **Center** category represents sources that strive for

[1] https://zenodo.org/record/4277311#.ZF0HIHZBxPY
[2] https://mediabiasfactcheck.com/

minimal bias, aiming to present information objectively. We have the **Left** category, which indicates media sources that are moderate to strongly biased in favor of liberal causes. The **Left-Center** category denotes sources that exhibit a slight to moderate bias towards liberal causes. On the other side of the spectrum, we have the **Right** category, representing media sources with a moderate to strong bias towards conservative causes. Finally, the **Right-Center** category includes sources that demonstrate a slight to moderate bias towards conservative causes. Table 1 includes a selection of news media outlets that fall into these five categories.

We extracted the names of all media sources with a bias category value falling within the aforementioned categories.

## 3.2 Data Extraction

This section delves into the methodology of extracting data from the various data sources and the steps taken to create the final dataset which contains social movement-related tweets that were reported in news media.

We start our data extraction process from Wikidata. From Wikidata, using occupations as a filter, we collect the information of all the celebrities and politicians who have a Twitter handle. In addition to the Twitter handle, various other attributes were obtained from Wikidata. The details of these additional features will be elaborated upon in the following section. From the list of users in the celebrity user group, we retained only users with a minimum of 200 quotations attributed to them in Quotebank. Since we are primarily focused on U.S. politicians, we filtered for politicians with party affiliation as either Democrat or Republican. In total, we extracted the information of 2360 celebrities and 1972 U.S politicians from Wikidata. We found coverage of 66% when we compared our list of politicians to the U.S. senators and representatives who served from 2011 to 2021[3].

We leveraged the Twitter Academic API[4] to collect tweets posted by the Twitter accounts of celebrities and politicians whose Twitter handles were extracted from Wikidata. Specifically, we collected tweets posted between 2017 and 2020. We chose 2017 as the starting year for collecting tweets, as this was the year when #MeToo movement gained global attention. Additionally, the MFOL movement started in 2018. Since Quotebank only had speaker-attributed quotations until April 2020, we collected tweets until 2020, resulting in the extraction of approximately 17 million tweets. As discussed in the last section, to retrieve social movement-related tweets, we used a list of social movement-related hashtags. Using hashtags as a filter to collect tweets ensures a high-quality dataset. We tried to expand this dataset by fine-tuning a BERT [21] language model to predict if a tweet text is related to a social movement or not. The social movement-related tweets retrieved using hashtags were used as the training data. However, the model had a very low recall (33%). Hence we

[3] `https://bioguide.congress.gov/`
[4] `https://developer.twitter.com/en/products/twitter-api/academic-research`

decided to continue our experiments with the initial dataset created using hashtags as a filter. Table 2 shows the number of tweets we were able to extract using the relevant hashtags.

The first step of our data extraction also involved retrieving the QID of all users belonging to the two user groups from Wikidata. Since Quotebank also uses QID to identify users, we leveraged this information to extract quotations attributed to these users from the Quotebank dataset. For consistency, we filtered the Quotebank quotations to only include those from the years 2017 to 2020. After applying the aforementioned steps, a subset of Quotebank was obtained which comprises all the news quotes attributed to the users belonging to our user groups, during the selected time frame.

In the upcoming subsection, we will retrieve tweets that were mentioned as quotations in news articles by utilizing the quote dataset and the social movement tweets dataset.

**Table 2** Number of social movement-related tweets extracted from Twitter for each user group using the relevant hashtags.

| **User group** | **No.of MeToo tweets** | **No.of MFOL tweets** |
|---|---|---|
| Celebrities | 5540 | 9005 |
| Politicians | 4481 | 10774 |

**Table 3** Number of reported social movement-related tweets extracted from Quotebank for each user group using Jaccard similarity to measure the overlap between a tweet and a quote.

| **Social Movement** | **No.of tweets Celebrities** | **No.of tweets U.S Politicians** |
|---|---|---|
| MeToo | 180 | 68 |
| MFOL | 91 | 75 |

### 3.2.1 Measuring Diffusion from Tweets to Quotes

Having obtained a dataset of social movement-related tweets from Twitter and a set of news quotes attributed to the same users from Quotebank, we can now determine which of these tweets were reported in news media. This was accomplished by comparing the tweets posted by a user with the quotes attributed to that user in our quotation dataset.

In our study, we noticed that different news outlets utilized distinct techniques when quoting a tweet in their articles. We identified the following three patterns that were commonly employed:

- The article used the entire tweet as a direct quotation.
- A tweet is broken down into multiple quotations and reported in the article. In some cases, only a portion of the tweet got reported (sub-tweet).
- A single quotation is created by combining one or more tweets from the same author.

Because of the various reporting styles by news outlets, a tweet may appear in any of the three forms described above. As a result, when looking for tweets reported by news outlets, we had to employ methods that measure the overlap between a tweet and a quote. We evaluate the four methods (discussed in section 2.1.2 of Chapter 2) listed below to accomplish this task:

- Least Common Sub-string (LCS)
- Jaccard similarity
- Levenstien distance
- Cosine similarity between sentence embeddings

All four methods used pre-processed tweet text. The first 3 methods work directly on the text content. For the last method, we computed the sentence embeddings for both tweet and quote text using sentence transformers [55]. To compare the efficacy of these methods, we normalized each method's similarity score to a value between 0 and 1. For each similarity method, we divided the tweet-quote pairs based on their similarity score into 10 equal-width bins (0.0 - 0.1, 0.1 - 0.2, ..). From each bin, we randomly selected 10 tweet-quote pairs and manually checked for accuracy. Figure 3 shows the plot between the accuracy of the methods. Cosine similarity and LCS were the least efficient and Jaccard similarity performed well with a 100% accuracy to the 0.8 similarity score. Hence we selected Jaccard similarity with a similarity score of 0.8 as the threshold to retrieve the tweets that were reported as quotes by news media. Table 3 details the amount of social movement-related quotes we could retrieve from Quotebank using this method.

At the end of the data extraction, we have two datasets that will be used during the experiments in Chapter 4.

- **Tweets Dataset**: Contains all the reported and NOT reported social movement-related tweets.
- **Articles Dataset**: Contains data from Quotebank for news articles that reported social movement-related tweets.

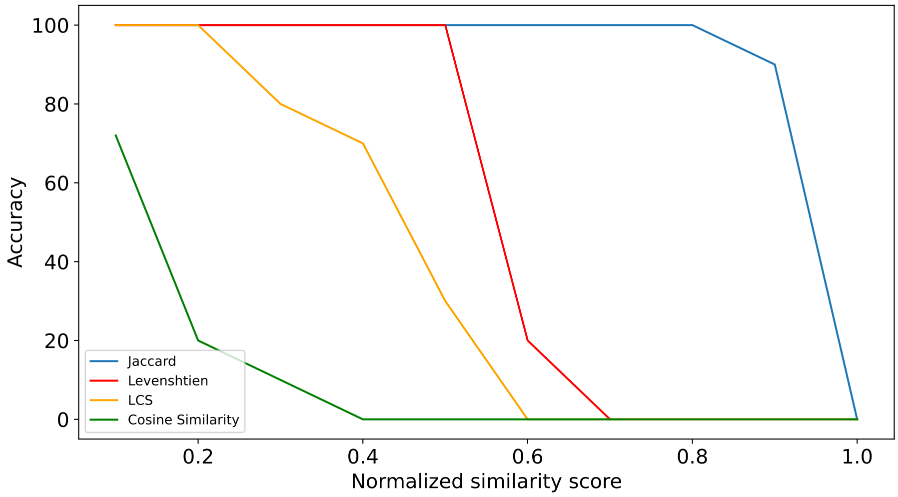


**Figure 3** Accuracy of the text similarity methods in measuring the overlap between tweets and quotes. In x-axis we have the normalized similarity score.

## 3.3 Data Features

In the preceding sections, we elaborated on the data sources and the data extraction process employed in this study. Due to the utilization of multiple data sources, we have access to a diverse set of data features, which we incorporate in our experiments. This section provides a detailed description of these features.

### 3.3.1 User-level Features

User-level features are unique to each user. We have extracted user-level attributes from Wikidata and Twitter. Table 9 provides a description of all the user-level attributes. The first three attributes in Table 9 were extracted from Wikidata, while the attributes *Followers*, *Following*, and *Total Tweets* were obtained from Twitter. The attribute *author involvement* is derived from other Twitter attributes.

**Table 4** Tweet-level attributes extracted from Twitter

| Feature | Description | Value range |
|---|---|---|
| Tweet text | Textual content of a tweet. | NA |
| Likes | Number of times a tweet has been liked.<br>This feature is a good indicator of public engagement. | [0 - 97713] |
| Replies | Number of times the tweet has been replied to.<br>by other Twitter accounts | [0 - 10606] |
| Retweets | A count of how many times the tweet has been retweeted. | [0 - 45942] |
| Quoted tweets | How many times the tweet was retweeted with a<br>new comment. | [0 - 2413] |
| Timestamp | Indicates the time at which the tweet was posted. | NA |

### 3.3.2 Tweet-level Features

Each tweet has its own set of tweet-level features, even if it is posted by the same user. All the tweet-level features extracted directly from Twitter using Twitter API are listed in Table 4. *Tweet text* is one such feature that contains the text message posted by the user. Since it contains text data, more features can be generated from it. For instance, Table 5 contains various features which were created from the tweet text using perspective API [5] (Chapter 2) from Google. Similarly, Table 6 contains features that we extracted from tweet text using TweetNLP package [15]. Apart from this, we have 2 other tweet-level features which were derived from other Twitter features. They are listed below:

- **Overall Relevance**: This is a tweet-level feature to capture the relevance of the social movement when the author posted their tweet. For each social movement-relevant tweet in our dataset, we check the day on which it was posted and then retrieve the number of social movement-related tweets posted on Twitter 5 days before and after.

- **Sub-group Relevance**: This feature is similar to the above one with the only difference that here we check the number of social movement-related tweets posted within the user group.

Some of the features we have extracted from Twitter, both at the tweet and author levels, are subject to change over time. For instance, the "number of followers" for an author and the "number of likes" for a tweet are examples of such time-variant features. Unfortunately, the Twitter API does not allow access to the historical values of these attributes. Therefore, we have to rely on the values available at the time of data extraction. However, it is worth noting that previous research [78] has demonstrated that a tweet tends to receive the majority of its retweets within the first hour of being posted. Consequently, for this study, we make the assumption that the tweet-level features, which closely capture public engagement, reflect the values observed by the news outlet when they picked up the tweet for reporting.

Apart from the tweet-level and user-level features discussed above, we also use the quotation-level features from Quotebank discussed in section 3.1.3. Additionally, we use the news-outlets bias category from section 3.1.4 in several of our experiments.

## 3.4 Preliminary Data Analysis

In this section, we share the results of the preliminary data analysis we conducted for this thesis.

[5] `https://perspectiveapi.com/`

**Table 5** Tweet-level attributes extracted from tweet text content using Perspective API

| Feature | Description | Value range |
|---|---|---|
| Severe toxicity | Tweet containing very hateful, aggressive, disrespectful text. | [0 - 1] |
| Toxicity | Tweet containing rude, disrespectful, or unreasonable text. | [0 - 1] |
| Insult | Tweet containing insulting, inflammatory, or negative content towards a person or a group of people. | [0 - 1] |
| Threat | Tweets which describe an intention to inflict pain, injury, or violence against an individual or group. | [0 - 1] |
| Identity attack | Tweet containing negative or hateful comments targeting someone because of their identity. | [0 - 1] |
| Profanity | Tweet text with swear words, curse words, or other obscene or profane language. | [0 - 1] |

**Table 6** Tweet-level attributes extracted from tweet text content using TweetNLP package [15]

| Feature | Description | Value range |
|---|---|---|
| Sentiment | Categorical feature which indicates the sentiment of the tweet. | Neutral, Positive, Negative |
| Hate Speech | A categorical feature which indicates if a tweet is hateful towards someone or a group of people. | hate, non-hate |
| Emotion | Emotion associated with a tweet. | angry, joy, sadness, optimism |
| Offensive Language | Identifies if a tweet contains offensive language. | offensive, non-offensive |

### 3.4.1 Tweets Reported as News Quotes

Table 3 presents the number of tweets related to social movements retrieved from the Quotebank dataset. Similarly, Table 2 provides the number of tweets related to social media posted by both user groups. In general, we observed a low level of content diffusion from social movement-related tweets to news quotes. For the #MeToo movement, only 0.032% of tweets posted by Celebrities were reported as news quotes, while 0.0075% of tweets posted by Politicians were reported. Likewise, for the #MFOL movement, 0.02% of tweets posted by Celebrities were reported as news quotes and 0.0069% of tweets posted by Politicians were reported. In both social movements, the Celebrity user group had a higher proportion of their social movement tweets reported in news media compared to Politicians.

#### Reported Tweet Statistics by Year

We examined the number of reported tweets for each year in both social movements. As mentioned earlier, the #MeToo movement began in 2017, while the

#MFOL movement started in 2018. Table 7 provides a breakdown of the reported tweets for all combinations of social movement user groups. Upon reviewing Table 7, we observed that the majority of reported tweets for the #MeToo movement were posted in 2017 and 2018. Therefore, we decided to focus solely on the #MeToo-related tweets from 2017 and 2018 for our study. Similarly, for the #MFOL movement, we are only considering the tweets posted in 2018 and 2019.

**Table 7** Number of social movement-related tweets reported by news media outlets each year

| **Social Movement-User group** | **2017** | **2018** | **2019** | **2020** |
|---|---|---|---|---|
| MeToo-Celebrity | 26 | 133 | 16 | 7 |
| MeToo-Politician | 0 | 58 | 9 | 0 |
| MFOL-Celebrity | 0 | 79 | 9 | 3 |
| MFOL-Politician | 0 | 53 | 22 | 0 |

**Table 8** Availability of Wikidata attributes in % for each user group. The "Field of work" attribute is not applicable (NA) for Politicians since all the users in this group have the same value for this attribute.

| **Wikidata attributes** | **Availability in % Celebrities** | **Availability in % U.S Politicians** |
|---|---|---|
| Occupation | 100 | 100 |
| Gender | 100 | 100 |
| Nationality | 99.8 | 98.07 |
| Date of Birth | 100 | 93.50 |
| Languages spoken | 100 | 100 |
| Political party membership | 15.4 | 100 |
| Religion | 20.5 | 25 |
| Field of work | 22.7 | N.A |
| Ethnic group | 12.3 | 4.6 |
| Position held | 5.02 | 87.06 |
| Academic degree | 1.11 | 4.5 |
| Candidacy | 1.67 | 6.8 |

### Sub-Tweets

As discussed in section 3.2.1, various news outlets employed different techniques when including tweets in their articles. Some outlets divided a tweet into multiple quotations, while others only reported a portion of the tweet, which we refer to as sub-tweets. During our manual analysis of these sub-tweets, we noticed that certain news outlets omitted hashtags and mentions contained within the tweet. To address this, we excluded tweets that had non-reported sub-tweets with less than

10 characters. From our observations, 371 news articles reported only a portion of the original tweet. We will use this subset of the original dataset to analyze the emotional tone of the reported sub-tweets. This analysis aims to determine whether news media prefer to report tweets with a negative emotion (discussed in section 4.2 of Chapter 4).

**Table 9** User-level Attributes extracted from Wikidata and Twitter

| **Feature** | **Description** | **Value range** |
|---|---|---|
| Occupation | This is the occupation value maintained in Wikidata. | Actors, Directors, Producers, Writers |
| Political party membership | This feature is not considered for celebrities as most of the users do not have this entry in Wikidata. | Democrat, Republican |
| Gender | Gender of the user as maintained in Wikidata | Female, Male, non-binary, TS female, TS male, CS female, genderfluid |
| Followers | Total number of Twitter accounts which follow the Twitter user | [1,690 - 108,880,185] |
| Following | Total number of Twitter accounts followed by the Twitter user | [0 - 1344714] |
| Total tweets | Total number of tweets posted by the Twitter user. This also indicates the user's engagement with the platform. | [108 - 387572] |
| Author Involvement | This feature captures the author's involvement in the social movement. It is a user-level feature that is calculated by dividing the total number of social movement-relevant tweets posted by a user by the total number of social movement-relevant tweets within the user group. | [0 - 1] |

### Modified Tweets

In our analysis, we discovered instances where news outlets made modifications to the original content of tweets posted by users. We identified 274 news articles that exhibited such modifications during our study. Further examination revealed that the majority of such articles added additional words, such as names of individuals and locations, to provide readers with more contextual information. Additionally, some articles altered the original tweet content by substituting profanity with placeholder strings like 'bleep' or 'f***'. However, a more comprehensive investigation into this subject is beyond the scope of this thesis, and therefore, we will not delve further into this aspect.

### 3.4.2 Available Wikidata Features

As discussed before in this chapter, Wikidata provides multiple user-level attributes for each user. However, we cannot use all such attributes because not all the users in our user group may not have these attributes available on their wiki data page. Table 8 gives a summary of the percentage of users in our user groups having entries in the Wikidata for a specific feature. Based on the results from Table 8, we picked the following user-level features from Wikidata: *Occupation*, *Political Party membership*, and *Gender*.

CHAPTER 4

# Methods, Analysis, and Results

In the current information age, social media has become a significant platform for social movements, replacing physical demonstrations to an extent. Previous studies [3, 19] suggest that both the public perception of social movements and even the decision-making of governing authorities can be influenced by the attention and coverage of news media. Therefore, this thesis aims to investigate the diffusion of social media content to news media in the context of social movements, by addressing the research questions and hypotheses discussed in the below sections.

## 4.1 How Tweets are Selected for News Reporting

News Gatekeeping is a crucial process in modern public life, whereby a vast amount of information is curated and condensed into a limited number of messages that are disseminated to the public. The gatekeeping process not only involves the selection of information but also influences the content and nature of the messages that are disseminated as news. This process, therefore, plays a critical role in shaping public opinion and perception on various issues [62].

The below research question aims to investigate the role of gatekeeping in the news selection process pertaining to the publication of tweets related to social movements. The focus is on understanding how gatekeeping shapes the dissemination of social media content to the news media, and its implications for public discourse. Unlike other types of news content, tweets possess unique author-level and tweet-level attributes that are associated with them (discussed in section 3.3 of Chapter 3). The study hypothesizes that these attributes can shed light on the gatekeeping process, and aid in understanding the factors that influence the selection of tweets for publication in the news media.

**Research Question 1 (RQ1):**
What is the relative influence of tweet-level and author-level attributes on the gatekeeping process for the selection of tweets related to social movements in the news media, and which of these attributes has the greatest impact on the likelihood of a tweet being published?

***Hypothesis 1.1 (H1.1):*** Tweet-level attributes that capture public engagement such as the number of likes, and number of replies have a greater influence on the gatekeeping process than other attributes.

***Hypothesis 1.2 (H1.2):*** The gatekeeping process employed by news media for the selection and publication of tweets differs for tweets posted by celebrities as compared to politicians.

### 4.1.1 Investigative Method

To test the hypothesis we defined above, we utilized the dataset comprising all social movement-related tweets from both user groups (**Tweets Dataset**). We divided this dataset into four subsets: MeToo-Politician, MeToo-Celebrity, MFOL-Politician, and MFOL-Celebrity to explore the differences in news selection processes for different social movements among authors in the two user groups. Each tweet in the dataset was supplemented with the author-level attributes of the tweet's author. For our study, we classified tweets into two groups: the treatment group ($treatment = 1$), which represents "tweets reported by the news media," and the control group ($treatment = 0$), which represents "tweets not reported by the news media." To analyze the impact of tweet-level and author-level attributes on the news selection process, we conducted a logistic regression analysis with all tweet and author attributes as independent variables and treatment as the dependent variable. The objective of this analysis was to identify how much of the variation in treatment could be explained by the independent variables.

### 4.1.2 Logistic Regression Analysis

We employed a logistic regression model to investigate the reporting behavior of news media toward social movement-relevant tweets posted by users in our user group. Our analysis incorporated various author-level and tweet-level features. For each social movement relevant tweet $\mathbf{t}$ posted by an author $\mathbf{a}$, let $rt[t]$ be the number of retweets $\mathbf{t}$ received, similarly $rp[t]$ be replies, $lk[t]$ be the likes, $qt[t]$ be the quoted tweets, $fc[a]$ be followers the author of tweet have, $fg[a]$ be the following, $tc[a]$ be the total tweets, $or[t]$ be overall relevance, $sr[t]$ be sub-group relevance, $ai[a]$ be the author involvement, $G[a]$ be the gender of author, $O[a]$ be the author's occupation, $ST[t]$ be tweet sentiment, $HS[t]$ be the hate speech, $EM[t]$ be the associated emotion, $OF[t]$ be the offensive language, $st[t]$ be sever toxicity, $to[t]$ be toxicity, $th[t]$ be threat, $in[t]$ be insult, $ia[t]$ be identity attack and $pr[t]$ be profanity. More descriptive details about all tweet level and author level attributes are present in Tables 9, 4, 5, and 6 in Chapter 3. We then define the model (in r-notation) as,

$$
\begin{aligned}
r = {} & rt[t] + rp[t] + lk[t] + qt[t] + fc[a] + fg[a] + tc[a] + or[t] \\
& + sr[t] + ai[a] + ST[t] + HS[t] + EM[t] + OF[t] + O[a] + G[a] \\
& + st[t] + to[t] + th[t] + in[t] + ia[t] + pr[t]
\end{aligned}
$$

where the dependent variable **r** equals 1 if the tweet **t** got reported and 0 otherwise. Here $ST$, $HS$, $EM$, $OF$, $G$, $O$ are categorical variables. Before training the logistic regression model, all the continuous predictors were standardized. To estimate the model parameters, we utilized the Broyden–Fletcher–Goldfarb–Shanno (BFGS) optimization method.

As this is an observational study, the selection of tweets reported by news media is non-random and treatment assignment is not based on randomization. This makes it difficult to compare the treatment group (reported tweets) with the control group (unreported tweets). To make the two groups comparable, we performed propensity score matching (discussed in section 2.1.3.2 Chapter 2) by considering all author and tweet-level features. Table 10 provides the number of tweets available in the treatment and control groups for each of the four social movement-user group combinations.

**Table 10** Number of tweets in treatment and control group before and after propensity matching

| Social Movement - User Group | Treatment | Control (Before Matching) | Control (After Matching) |
|---|---|---|---|
| MeToo - Celebrity | 159 | 4500 | 151 |
| MeToo - Politician | 67 | 4413 | 64 |
| MFOL - Celebrity | 78 | 6499 | 76 |
| MFOL - Politician | 75 | 9866 | 72 |

We performed the Wilcoxon signed-rank test (discussed in section 2.1.3.3 of Chapter 2) on all predictors in the matched dataset. The summary of the results for both social movements and user groups is presented in Table 11. Statistically significant predictors ($p<0.05$) are marked in bold.

We identified four tweet-level attributes that measure public engagement for a tweet[1], namely the number of likes, retweets, quoted tweets, and replies. As presented in Table 11, our analysis reveals that at least one of these public engagement features is statistically significant for all social movement-user group combinations. For the statistically significant features identified, we calculated the mean values for both the treatment and control groups. Our analysis revealed that the mean values for these features were consistently higher in the treatment group than in the control group, providing support for our hypothesis **H1.1** that tweet-level attributes which

[1]https://developer.twitter.com/en/docs/twitter-api/metrics

**Table 11** P-values of the predictors after performing the Wilcoxon signed rank test on tweets after matching using propensity score. Predictors with significant p-values ($p < 0.05$) are marked in bold.

| **Predictor** | **P-Value MeToo-Celeb.** | **P-Value MeToo-Polt.** | **P-Value MFOL-Celeb.** | **P-Value MFOL-Polt.** |
|---|---|---|---|---|
| No.of retweets | 0.549 | 0.663 | 0.395 | 0.785 |
| **No.of replies** | **0.0045** | 0.066 | **0.0017** | **0.000012** |
| **No.of likes** | **0.0130** | 0.672 | **0.0022** | 0.161 |
| **No.of quoted tweets** | **0.0138** | **0.0141** | **0.00046** | **0.00042** |
| gender | 0.122 | 0.654 | 0.188 | 0.479 |
| occupation | 0.949 | - | - | - |
| no.of followers | 0.5670 | 0.857 | 0.798 | **0.015** |
| no.of accounts following | 0.5059 | **0.045** | 0.781 | 0.7922 |
| total tweets posted | 0.6708 | 0.6414 | 0.888 | **0.035** |
| overall movement relevance | 0.2946 | **0.019** | 0.1811 | 0.1908 |
| sub-group movement relevance | 0.5722 | 0.758 | 0.1820 | 0.147 |
| author involvement= | 0.6708 | 0.797 | 0.114 | **0.019** |
| Tweet sentiment | 0.5109 | 0.11 | 0.459 | 0.948 |
| Hate speech | 0.1572 | 0.317 | - | 0.563 |
| Emotion | 0.3218 | 0.342 | 0.915 | 0.6925 |
| Offensive Lang. | **0.0423** | 1.0 | 0.654 | 0.317 |
| Severe toxicity | 0.2112 | 0.287 | 0.832 | 0.673 |
| toxicity | 0.3315 | 0.203 | 0.857 | 0.340 |
| threat | 0.8980 | 0.203 | 0.966 | 0.428 |
| insult | 0.3248 | 0.250 | 0.920 | 0.360 |
| **identity attack** | 0.5628 | **0.026** | 0.450 | 0.787 |
| profanity | 0.4713 | 0.419 | 0.808 | 0.627 |

capture public engagement have a greater influence on the gatekeeping process than other attributes.

To test the validity of our second hypothesis (**H1.2**), we examined the statistically significant features in the matched dataset for both the Politicians and Celebrity user groups. The results indicate that for the Celebrity user group, the *number of likes* was found to be a significant predictor for getting a tweet reported in both social movement tweets. On the other hand, the *number of likes* feature did not appear to be a significant predictor for the Politicians user group in both social movements. This could indicate that audience preference is taken into account while reporting tweets posted by celebrities, but a similar process is not followed while reporting tweets posted by Politicians. This provides support for our hypothesis **H1.2** that there is a difference in the gatekeeping process employed by news media for the selection and publication of tweets posted by celebrities as compared to politicians.

### 4.1.3 Classifiers

To corroborate our results, we trained four machine learning models (section 2.1.4 of Chapter 2) for the binary classification task of predicting whether a tweet gets reported by news media. All the tweet and author-level features were included as input during the model training. The best-performing model had the F1-Score of 68% in predicting whether a tweet gets reported. Additionally, we examined the most important features identified by this model and found that tweet-level attributes related to public engagement, such as the number of likes, retweets, and replies, were given more importance by the model. The performance of all the models we trained is provided in appendix Table 15.

## 4.2 Negative News Coverage in News Media

An observation made during the initial data analysis in Chapter 3 was that news outlets tend to selectively report portions of the original tweet text (sub-tweet) as direct quotations. Previous studies [4, 34, 61] have reported an over-representation of negative news coverage pertaining to the economy and politics in various news media outlets and the audience exhibits a preference for negative news content as well [68]. Below research question aims to investigate if news outlets exhibit consistent behavior as above when they are selectively reporting portions of a given tweet.

**Research Question 2 (RQ2):**
To what extent do news outlets exhibit a preference for reporting sub-tweets that convey negative emotions expressed by the author?

***Hypothesis 2.1 (H2.1):*** News outlets tend to favor sub-tweets that capture negative author emotions because such tweets are more likely to attract reader attention and generate engagement.

### 4.2.1 Investigative Method

As stated earlier our investigation focuses on the tweets that were only partially reported by news outlets. For tweets with only portions of the original content reported by the news media, we define two groups: the treatment group ($treatment = 1$) represents "reported sub-tweet text content," and the control group ($treatment = 0$) represents "unreported sub-tweet text content." To determine the author emotion of the sub-tweet content, we utilize the four emotion detection models discussed in Chapter 2 to obtain emotion labels, creating four distinct dataset versions. We then perform our experiments for each dataset version to investigate whether the news media preferentially reports sub-tweet text content with negative emotions. We use a logistic regression model with the sub-tweet emotion label as the independent variable and treatment as the dependent variable.

### 4.2.2 Logistic Regression Analysis

We used a logistic regression model to investigate whether the news outlets tend to report sub-tweets that capture negative author emotion (**RQ2**). Our analysis incorporated the emotion of the sub-tweet text content extracted using the emotion detection models we discussed in Chapter 2. For sub-tweet content $\mathbf{t}$, let $e[t]$ be the author emotion. We then define our model as,

$$r = e[t]$$

where the dependent variable $r$ equals 1 if the sub-tweet was reported by the news outlet and 0 otherwise.

In the below sub-sections, we discuss the results when different models are used to get the author emotion.

#### Model 1

Model 1 [2] uses a combination of BERT [21] and stacked Bi-LSTM [26] architecture. And it is fine-tuned on the GoEmotions dataset [20]. During our analysis, we found emotions **anger**, **love**, and **realization** to be statistically significant negative predictors in a sub-tweet getting reported by news outlets.

#### Model 2

Model 2 [54] uses a BERT-BASE [21] model fine-tuned on EMO dataset [17]. During our analysis, we found emotions **sadness** as a statistically significant positive predictor for a sub-tweet to be reported.

#### Model 3

Model 3 [32] uses a DistilRoBERTa-base model fine-tuned on 6 different emotion detection datasets. During our analysis, we found emotions **disgust** as a statistically significant positive predictor for a sub-tweet to be reported. And we found emotions **fear**, **joy**, **neutral**, **sadness**, and **surprise** to be statistically significant negative predictors. Error plot for the same is shown in the figure 4.

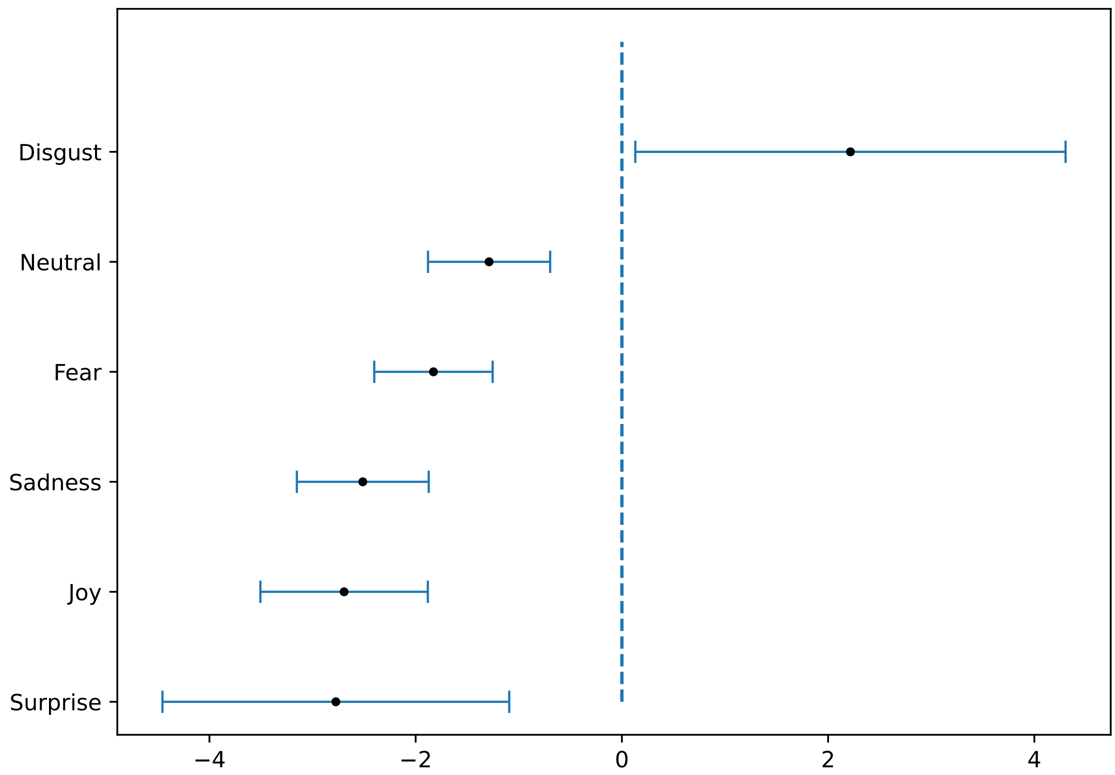


**Figure 4** Error plot illustrating positive and negative predictors in logistic regression analysis conducted using emotion labels from Model 3. All the emotion labels in the figure are statistically significant.

**Model 4**

Model 4 is based on LIWC words [52]. Here to detect the emotion of the sub-tweet content, we count the number of positive emotion words and negative emotion words contained in a sub-tweet and take their difference. In our analysis, we found that the emotion **neutral** is a statistically significant positive predictor for a sub-tweet to be reported. We have summarized the results from all four models in Table 12.

We could observe that the emotions which depict positive emotions such as love, joy, and surprise are often negative predictors of getting a tweet reported. And negative emotions such as disgust and sadness are positive predictors.

## 4.3 Author Selection Bias

As outlined in Chapter 3, we opted to use the author user groups of Celebrity and US Politicians in our experiments, based on the fact that news outlets exhibit a preference for covering stories related to high-profile individuals and individuals in

**Table 12** Summary of all statistically significant positive and negative predictors in logistic regression analysis conducted using the 4 emotion detection models.

| **Model** | **Statistically significant positive predictor** | **Statistically significant negative predictor** |
|---|---|---|
| Model 1 | - | anger, love, realization |
| Model 2 | sadness | - |
| Model 3 | disgust | fear, joy, neutral, sadness, surprise |
| Model 4 | neutral | - |

positions of power [31]. Nonetheless, it is important to note that the authors belonging to these user groups exhibit differences among themselves and possess distinct author-level characteristics, including gender, occupation, and others. Moreover, there exist significant differences in the popularity levels of these authors. For instance, more renowned authors tend to have a higher number of Twitter followers, which could potentially result in higher public engagement for their tweets. Taking into account all of these factors, our subsequent research question seeks to investigate whether news outlets exhibit any discernible bias in their selection of which authors to report.

**Research Question 3 (RQ3):**
To what extent do news outlets exhibit bias in their selection of authors to report on tweets in the context of social movements?

***Hypothesis 3.1 (H3.1):*** News outlets demonstrate a bias by exhibiting a greater tendency to report on tweets related to the #MeToo movement that have been authored by women.

***Hypothesis 3.2 (H3.2):*** News outlets exhibit a bias in their reporting on tweets related to social movements by selectively featuring those authored by highly popular authors.

### 4.3.1 Investigative Method

To test the hypothesis outlined earlier, we utilize a dataset that includes the author-level attributes of all users who posted at least one social movement-related tweet. For our study, we classified authors into two groups: the treatment group ($treatment = 1$), which represents "authors who were reported by the news media," and the control group ($treatment = 0$), which represents "authors who were not reported by the news media."To analyze the impact of author-level attributes on the selection of authors to quote, we conducted a logistic regression analysis with all author-level attributes as independent variables and treatment as the dependent variable. The objective of this analysis was to identify how much of the variation in treatment could be explained by the independent variables.

### 4.3.2 Logistic Regression Analysis

For this experiment, we utilized a logistic regression model to investigate the potential biases of news outlets in selecting which tweet authors to report. Our analysis included several author-level features that were introduced in Chapter 3. For each author **a** who posted at least one social movement-related tweet let $fc[a]$ be followers the author of the tweet has, $fg[a]$ be the following, $tc[a]$ be the total tweets, $G[a]$ be the gender of the author, $O[a]$ be the author's occupation. We then define the model in r-notation as,

$$r = fc[a] + fg[a] + tc[a] + O[a] + G[a]$$

where the dependent variable **r** equals 1 if the author **a** was reported and 0 otherwise. In our experiments with the Politician's user group, we excluded Occupation as an independent variable since all users shared the same occupation. Prior to training the logistic regression model, we standardized all continuous predictors. The BFGS optimization method was employed to estimate the model parameters.

Due to the non-randomized selection of authors for reporting, conducting a comparison between the treatment and control groups becomes challenging in this observational study. To address this issue, we performed propensity score matching by considering all the aforementioned independent variables.

We performed the Wilcoxon signed-rank test on all predictors in the matched dataset. The summary of the results for all the social movement user group combinations is available in Table 13.

After analyzing the results, we did not find any evidence to support the idea that news outlets exhibit bias in reporting #MeToo movement tweets posted by women. Therefore, we reject our hypothesis **H3.1**. In the analysis of the results of the Celebrity user group's tweets related to the MeToo movement, we found that the *number of followers* attribute is statistically significant, but the effect size is minimal. However, we did not observe similar effects for other user groups and social movements, leading us to reject our second hypothesis **H3.2** for this experiment.

## 4.4 Trend in Reporting Social Movements

Previous research has indicated that partisanship and ideology play a significant role in determining a person's support for social movements. For instance, a study found that while 85% of Democrats reported that the #MeToo movement increased their awareness of sexual harassment, only around 50% of Republicans expressed the

**Table 13** P-values of the predictors after performing the Wilcoxon signed rank test on authors after matching using propensity score. Predictors with significant p-values ($p < 0.05$) marked in bold.

| Predictor | P-Value MeToo-Celeb. | P-Value MeToo-Polt. | P-Value MFOL-Celeb. | P-Value MFOL-Polt. |
|---|---|---|---|---|
| Followers | **0.0413** | 0.279 | 0.256 | 0.238 |
| Following | 0.8138 | 0.433 | 0.388 | 0.281 |
| Total tweets | **0.0413** | 0.770 | 0.360 | 0.212 |
| Gender | 0.058 | - | 0.834 | 0.317 |

same sentiment [16]. Similar trends have been observed in the gun violence social movement, as previous research has shown that the impact of mass shootings on gun control legislation is influenced by the political party in power and in states where Republicans control the legislature, the number of laws that ease restrictions on guns increases twofold in the year following a mass shooting, as per the yearly statistics [45]. Given that news outlets have political biases (Chapter 3), we aim to examine how their reporting of social movement-related tweets may vary. This will be explored through our below research question.

**Research Question 4 (RQ4):**
How did the political bias of news outlets influence the reporting of social movement-related tweets?

Drawing upon the aforementioned discussion, which indicates that Democrats tend to be more supportive of the #MeToo movement than Republicans, our research seeks to investigate the following hypothesis:

***Hypothesis 4.1 (H4.1):*** News outlets with a left-leaning bias are more likely to report on the #MeToo movement compared to news outlets with other political biases.

***Hypothesis 4.2 (H4.2):*** News outlets with a right-leaning political bias will show a delayed start in reporting on social movement-related tweets compared to outlets with other political biases.

### 4.4.1 Investigative Method

To investigate the reporting patterns of news outlets with varying political affiliations, we utilized the dataset extracted from Quotebank's article-centric version, containing social movement-related tweets reported by news media (**Articles Dataset** from Chapter 3). Entries without an entry in the mediabiasfactcheck website were removed from this dataset. We then grouped the remaining entries by the bias label

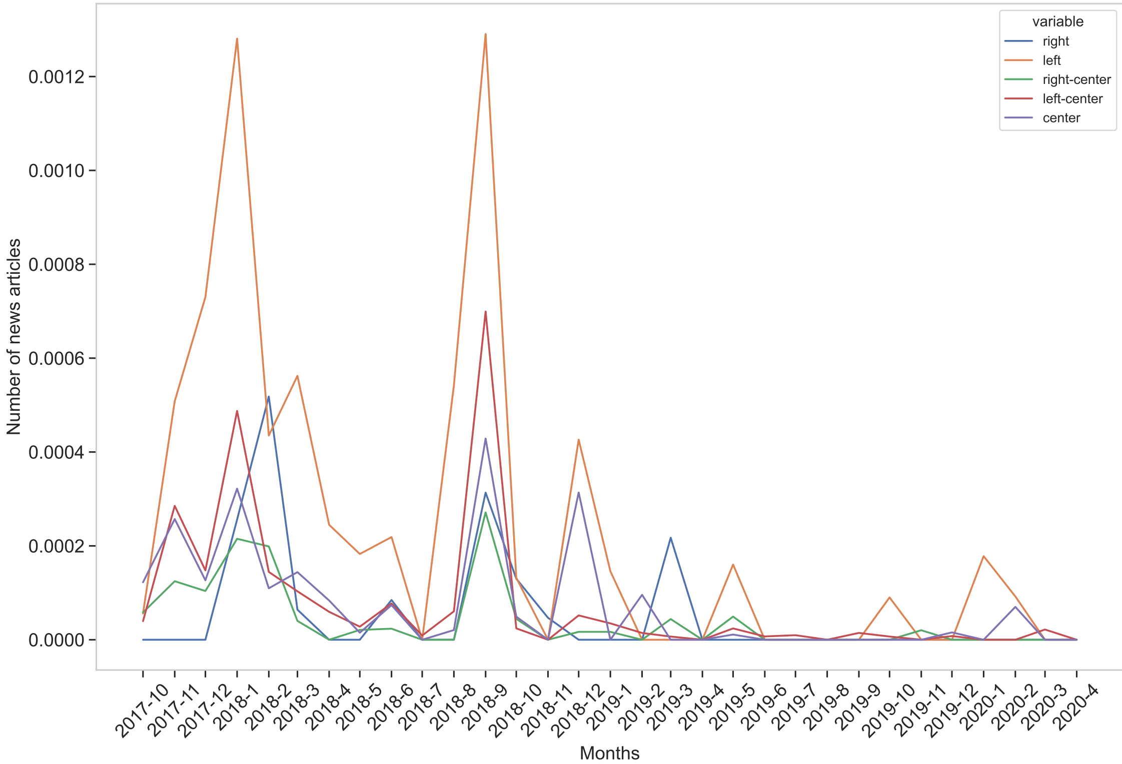


**Figure 5** Normalized time series plot for the monthly reported #MeToo related tweets posted by the users in the Celebrity user group.

(Chapter 3) of the respective news outlets. Monthly time series were obtained by calculating the total number of social movement-related articles published by news outlets in each bias category group each month, which were then normalized by dividing the count with the total number of articles published by these news outlets in the same months. A one-way ANOVA test (Chapter 2) was performed to determine any significant differences between these time series.

### 4.4.2 Analysis

Normalized monthly time series are plotted in figures 5, 6, 7 and 8.

To examine the variation in reporting of the #MeToo movement across news outlets with different polarities, we conducted a one-way ANOVA test. The number of #MeToo-related news articles published per month was the dependent variable, and the news outlet polarity was the independent variable. The results of the test revealed a p-value of 0.00035, indicating that at least one of the media bias group significantly differs from the overall mean of the dependent variable. To further investigate this difference, we conducted a posthoc Tukey HSD test [1]. The results of the test are shown in Table 14. The results indicate that news outlets with left bias have a significant difference in reporting #MeToo movement-related tweets when

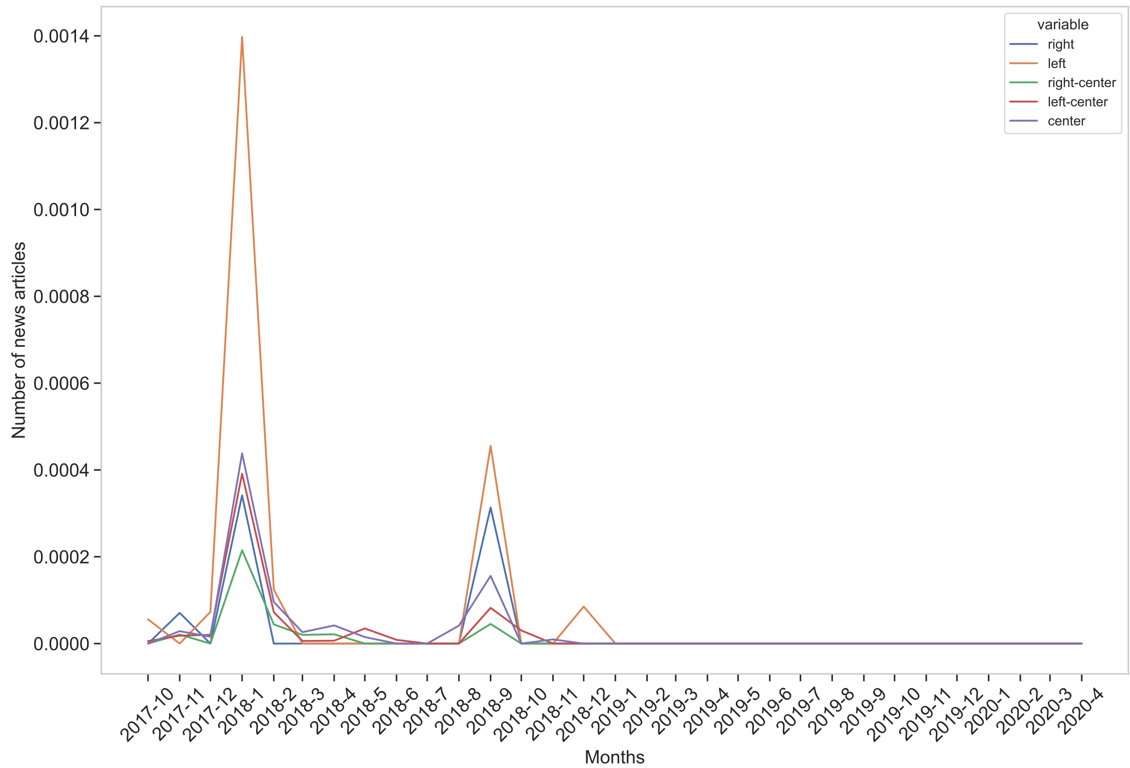


**Figure 6** Normalized time series plot for the monthly reported #MeToo related tweets posted by the users in the Politician user group.

compared to other news outlet bias categories. Thus we could verify the hypothesis **H4.1**. We conducted both tests for other social movements as well. However, we couldn't find any similar effects.

To verify the validity of the second hypothesis (**H4.2**), we take a close look at the time series plots. For the time series plot for reported #MeToo movement tweets posted by Celebrities, right-wing outlets did not report any tweets for the first 3 months of the movement (October 2017 - December 2017). However, we couldn't see a similar trend for other social movements and user groups. Hence we can only partially accept our second hypothesis **H4.2**.

## 4.5 Objectivity in Reporting

As mentioned earlier in this chapter, the news media's attention and coverage of social movements can shape public perception of them. For many, news articles are the primary source of information on current events. However, the news outlets can exhibit an inherent bias that is reflected in their reporting and this could impact the public perception as well [28]. Hence to comprehensively understand how news outlets report social movement-related tweets, it is crucial to investigate how the tweet authors are being portrayed in news articles.

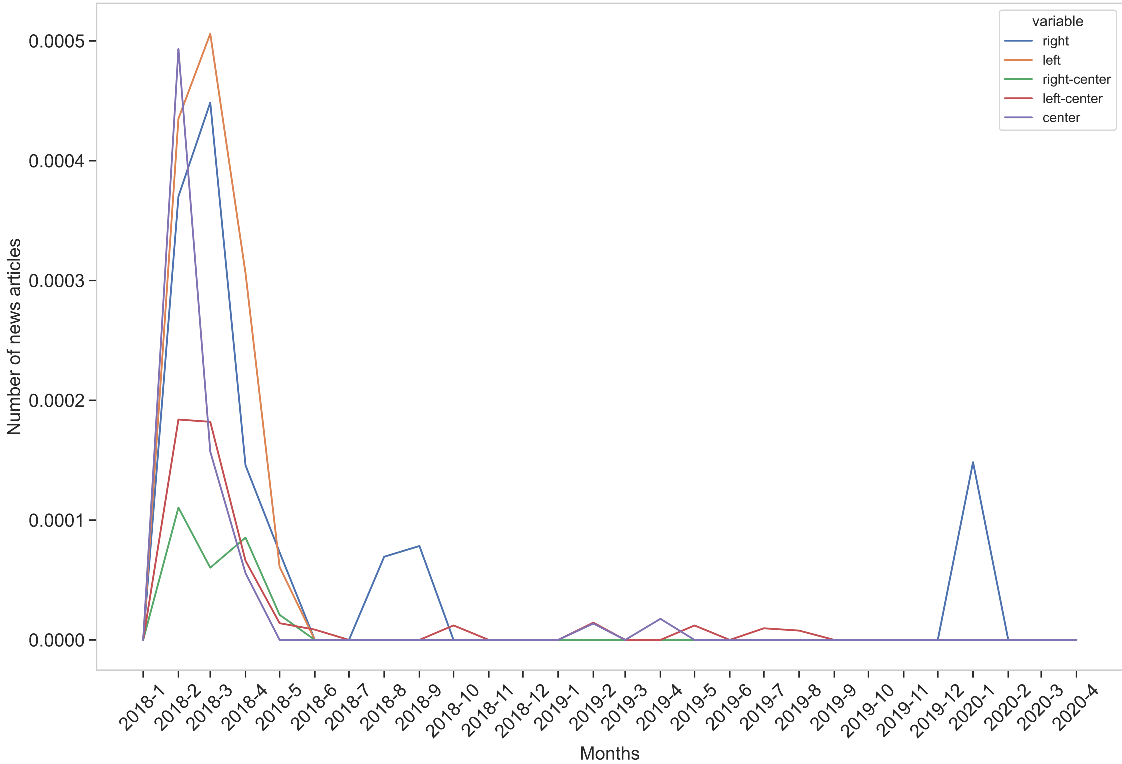


**Figure 7** Normalized time series plot for the monthly reported #MFOL related tweets posted by the users in the Celebrity user group.

**Research Question 5 (RQ5):**
How do news outlets portray the authors of social movement-related tweets they publish?

***Hypothesis 5.1 (H5.1):*** There is a higher likelihood of partisan news outlets portraying politicians from opposing political parties in a negative light.

### 4.5.1 Investigative Method

To investigate the hypothesis, we utilized the article-centric version of the dataset and filtered out news outlets without entries in medibiasfactcheck. The left-context and right-context text data from Quotebank were combined to obtain the news outlet's sentiment towards the author, which was then processed through our TSC pipeline as described in section 2.1.1.5 of Chapter 2. To examine how partisan news outlets depict politicians from opposing parties, we employed a logistic regression model with the objectivity of the news outlet in reporting the author as the dependent variable and news outlet bias and the author's party affiliation as independent

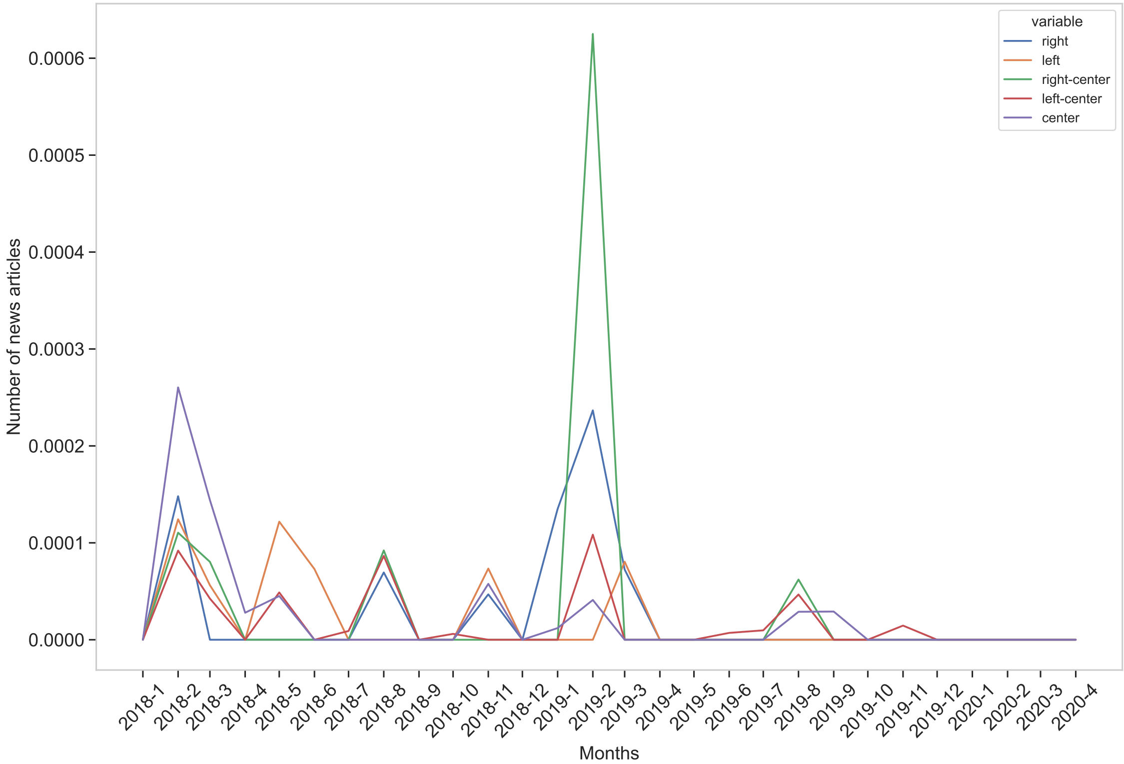


**Figure 8** Normalized time series plot for the monthly reported #MFOL related tweets posted by the users in the Politician user group.

variables. A positive or negative sentiment towards the author was assumed to indicate a lack of objectivity in the news outlet's reporting.

### 4.5.2 Logistic Regression Analysis

For this experiment, a negative or positive sentiment towards the author indicates a lack of objectivity in the news outlets reporting. Let $p[a]$ be the party affiliation of the politician **a** and $b[n]$ be the bias of the news outlet **n**, we define our model as,

$$y = p[a] + b[n] + (p[a] * b[n])$$

where **y** is the dependent variable which indicates the objectivity in reporting the politician.

$y = 0$ indicates politician was not objectively reported by the news outlet (positive/neutral sentiment)

$y = 1$ indicates the politician was objectively reported by the news outlet (neutral sentiment)

The term $p[a] * b[n]$ is used in the model to capture the interaction between the party affiliation of the politician and the bias of the news outlet.

**Table 14** Tukey HSD results for multiple time series comparisons. The table presents the pairwise comparisons of all time series and significant p- values ($p < 0.05$) marked in bold.

| **Group 1 New outlet Bias** | **Group 2 New outlet Bias** | **P-Value** |
|---|---|---|
| Center | Left | **0.0077** |
| Center | Left-Center | 1.0 |
| Center | Right | 0.9931 |
| Center | Right-Center | 0.9597 |
| Left | Left-Center | **0.0097** |
| Left | Right | **0.0018** |
| Left | Right-Center | **0.0007** |
| Left-Center | Right | 0.9875 |
| Left-Center | Right-Center | 0.9426 |
| Right | Right-Center | 0.9999 |

In this experiment, we couldn't find any evidence which could prove that partisan news outlets negatively portray politicians from opposing political parties. Hence we reject the hypothesis **H5.1**.

CHAPTER 5

# Discussion

In this chapter, we provide a detailed discussion of our results and their implications for each research question examined in the previous chapter. Additionally, we compare our findings with existing research and highlight our contributions.

## 5.1 Media Coverage of Social Movement Tweets

One of the primary focal points of this thesis revolves around the utilization of Twitter as a source of news by journalists. Previous studies [50, 14, 11] have conducted manual content analyses of news articles to identify reported tweets, focusing on various general topics such as politics and sports. In contrast, our work specifically focuses on tweets related to social movements. To assess the dissemination of these tweets, we retrieved social movement-related tweets from Twitter and employed the Jaccard similarity measure for identifying the tweets which were reported as direct quotations in news articles. This approach has allowed us to more accurately gauge the news coverage of social movement-related tweets as quotations within online news media. Overall, we discovered a relatively low volume of reported social movement-related tweets from both user groups (section 3.4.1 of Chapter 3). One possible explanation for this limited number of reported tweets is the variation in popularity among users within our user groups, both on Twitter and in society. Directly comparing our findings with previous studies [50, 14, 11] is challenging as the previous works solely concentrate on reported tweets and do not consider tweets that were not reported. To the best of our knowledge, this is the first study to examine the coverage of social movements in news media specifically with regard to the reporting of social movement tweets as news quotations.

## 5.2 How Tweets are Selected for News Reporting

Our study aims to explore the extent to which journalists take the public engagement of a tweet into account when reporting tweets related to social movements. When it comes to tweets, features like the number of likes and replies serve as effective indicators of the level of public engagement with the tweet. We analyzed all the social movement-related tweets posted by both user groups and found that tweet-level attributes that capture public engagement increased the likelihood of a tweet being reported. Based on our findings, it can be inferred that journalists take into account tweet characteristics that reflect public engagement when making decisions about whether or not to include a tweet in their article. Previous studies [8, 10, 73] have suggested that journalists and news organizations have become increasingly responsive to audience preferences. This could be a possible reason for our results. Additionally, our findings align with the research conducted by [50] who also found that public engagement cues, such as the number of likes and retweets, played a significant role in determining which tweets were included in news articles. However, it is important to note that [50] focused solely on tweets that were reported by news media, while our research takes into account both reported and non-reported tweets. Additionally, our study incorporates several other author and tweet-level features as independent variables, providing a more comprehensive analysis of the news selection process. Overall, our findings contribute to the understanding of how journalists consider the public engagement of tweets and various other attributes when deciding which social movement-related tweets to include in their news coverage.

To investigate the disparity in the gatekeeping process used by news media when selecting and publishing tweets from celebrities and politicians, we compared the statistically significant predictors for both user groups. One notable difference we identified was that the attribute "number of likes" served as a significant positive predictor for the celebrity user group, whereas it lacked statistical significance for the politician's user group. These findings suggest that journalists prioritize audience preference when reporting tweets from celebrities, while news organizations may take their own political bias into consideration when reporting tweets from politicians. Our results align with previous research [63] that found journalists embrace audience participation in creating "soft news" content but exhibit resistance when selecting political and other "hard news" content.

## 5.3 Negative News Coverage in News Media

In this study, our aim was to examine whether news outlets exhibited any bias in reporting sub-tweets that conveyed negative emotions expressed by the author. For this, we retrieved the author emotions of reported and non-reported sub-tweets using various models (discussed in Section 2.1.1.4). Our observations indicated that sub-tweets expressing positive emotions such as love, joy, and surprise tended to be less likely to be reported. Conversely, sub-tweets conveying negative emotions such as disgust and sadness were more likely to be positively associated with being reported. This is similar to the previous studies [4, 34, 61] which reported an over-representation of negative news in the media.

The results of this experiment exhibit subtle contradictions. For instance, when analyzing the emotion labels generated from model 3 in our experiments, we observed that sadness and fear were statistically significant negative predictors in getting a sub-tweet reported. However, when employing model 2, sadness emerged as a statistically significant positive predictor (refer to Table 12). It is worth noting that the emotion detection models utilized in our experiments encompass different numbers of target emotion labels. For example, model 1 encompasses 28 distinct target emotion classes, while model 4 focuses on only 3 target emotion classes. Consequently, these variations among the models can lead to different predictions of emotion labels for the same text content. Furthermore, some models possess a larger number of target emotion classes and demonstrate greater robustness in identifying sub-categories of emotions. Despite these disparities, we observed a notable trend in the preference of news outlets to report sub-tweets conveying negative emotions.

## 5.4 Author Selection Bias

To examine the bias displayed by news outlets in their selection of users to report on, we analyzed all users belonging to both user groups who had posted at least one tweet related to social movements. Given that the #MeToo movement predominantly addresses sexual harassment against women, we anticipated news outlets to exhibit a bias by featuring a higher number of tweets posted by women. Surprisingly, our experiments yielded no evidence of such an effect. We hypothesized that the inclusion of numerous users with lower popularity in our user list could potentially explain the absence of any significant results. Furthermore, we examined whether the author's popularity, as indicated by the number of followers, attracts more attention from journalists. In the analysis of the tweets from the Celebrity user group pertaining to the MeToo movement, we discovered that the attribute of follower count exhibited statistical significance, although the impact size was minimal. However, we did not observe similar effects for other user groups and social movements. Our findings align with a previous study [48] which also concluded that the popularity of Twitter accounts does not significantly influence the attention received from journalists. The absence of results may be attributed to the time-variant nature of the tweet-level feature "Number of Followers." As explained in Chapter 3, Twitter does not provide access to the historical values of any tweet-level features. Therefore, the number of followers feature does not precisely reflect the author's popularity at the time of posting a tweet related to a social movement. Repeating the experiment using a feature that accurately captures the author's popularity at the time of posting could potentially provide deeper insights into the actual impact of author popularity.

## 5.5 Trends in Reporting Social Movements

We explored the impact of political bias among news outlets on their coverage of tweets related to social movements. For this, we analyzed the monthly time-series data of the social movement-related tweets reported by outlets in each of the political bias categories. Our findings revealed that news outlets with a left-leaning bias exhibit a higher tendency to report tweets associated with the #MeToo movement. These results align with a previous study [39] that conducted a content analysis of 516 articles from four U.S. newspapers and five newspapers in South Korea, and observed that liberal media tends to report more on the #MeToo movement. Other studies [16] have also demonstrated that Democrats show greater support for the #MeToo movement compared to Republicans. Furthermore, during our experiments, we noted that the right-leaning outlets refrained from reporting any tweets associated with the #MeToo movement in the initial three months. We speculate that the political ideologies of news outlets contribute to this disparity in reporting.

## 5.6 Objectivity in Reporting

In our investigation, we examined the level of objectivity exhibited by news outlets when reporting on politicians who posted tweets related to social movements. We took into account both the political bias of the news outlet and the party affiliation of the politician. Our objective was to determine whether partisan news outlets had a greater tendency to present politicians from opposing political parties in a negative manner. However, our findings did not yield any evidence supporting the notion that partisan news outlets negatively portray politicians from opposing political parties. Our findings contradicted the results of a previous study [38], which indicated that news outlets frequently employ nonobjective quotatives when quoting politicians from opposing ideologies. It is worth noting that the aforementioned study had a significantly larger dataset compared to ours. Additionally, our research heavily relied on the limited news article text content available from Quotebank. Upon manual analysis of the outcomes generated by our best-performing model (discussed in Section 2.1.1.5), we observed an accuracy of only 54.3%. The absence of significant results can potentially be attributed to the scarcity of data and the quality of the available text content.

## 5.7 Limitations

We have identified several limitations that may have impacted the outcomes of our study. Firstly we discuss the limitations of the various data sources. Our reliance on Wikidata to compile the user lists for our two user groups introduces certain constraints. Users who are politicians or celebrities but lack a presence on Wikidata or have their Twitter handle information absent from Wikidata are excluded from our study. Consequently, we cannot assert the comprehensiveness of our user lists. For this study, we obtained all the tweets posted by users in both user groups spanning from 2017 to 2020. However, since our data collection took place in 2022, it is possible that some tweets may be missing from our dataset if the authors deleted their tweets or deactivated their Twitter accounts. Furthermore, in this research, we utilized various tweet-level features extracted from Twitter. As discussed in previous chapters, certain tweet-level attributes are subject to change over time, and extracting historical values for these attributes is not supported by Twitter's platform. Therefore, for our experiments, we relied on the values available at the time of extraction. In our study, our primary source for measuring the content diffusion of social movement-related tweets from Twitter to news media is Quotebank. It is important to note that Quotebank is generated using Spinn3r news data, which relies on the crawling of news media websites. As a result, there may be limitations in terms of comprehensive coverage of all news websites, potentially leading to gaps in overall coverage. Consequently, we cannot assert that our study captures the complete content diffusion of social media tweets due to these limitations in data collection.

Throughout our experiments, we have taken into account various attributes at both the tweet and author levels. However, it is important to acknowledge that in real-world scenarios, there are likely additional unobserved confounding factors that may influence the treatment of our experiments. For example, the decision to report a tweet may not solely depend on public engagement metrics. Other factors, including journalist biases and editorial decisions, could also play a significant role. Capturing such complex decisions within our experimental setup is challenging.

CHAPTER 6

# Conclusion

The main contribution of this thesis was to measure the content diffusion of social movement-related tweets to news quotes and explore the changes to the news gatekeeping process in this context. Our analysis revealed a relatively low volume of reported social movement-related tweets. We explored how journalists consider public engagement indicators such as likes and reply when deciding which social movement-related tweets to include in their articles. Our findings indicated that tweet characteristics reflecting public engagement increased the likelihood of a tweet being reported. Additionally, our results indicate that journalists prioritize audience preference while reporting celebrity tweets but consider their own political bias when reporting tweets from politicians. Our results suggest that journalists and news organizations pay attention to audience preference while reporting tweets and thus indicate the impact of social media and audience preference in the news gatekeeping process. We investigated the presence of negative news coverage in news media by analyzing sub-tweets conveying different emotions. Our observations indicated a preference for reporting sub-tweets expressing negative emotions, while those conveying positive emotions were less likely to be reported. However, the choice of emotion detection models led to some contradictory findings, highlighting the complexity of this analysis. To investigate author selection bias, we analyzed the selection of users to be reported on by news outlets. Surprisingly, we found no evidence of a bias favoring tweets posted by women in the context of the #MeToo movement. Additionally, the popularity of Twitter accounts did not significantly influence the attention received from journalists, consistent with a previous study. However, the absence of significant results may be attributed to the limited data available and the time-variant nature of certain tweet-level features. Examining trends in reporting social movements, we observed that news outlets with a left-leaning bias exhibited a higher tendency to report tweets associated with the #MeToo movement. This aligns with previous research and indicates a political ideology-based disparity in reporting. We also explored the level of objectivity in reporting social movement-related tweets posted by politicians. However, our findings did not support the notion that partisan news outlets negatively portray politicians from opposing political parties, although the limited data and text content quality may have influenced the results. Several limitations should be acknowledged, including constraints related to data sources, potential gaps in data coverage, and the challenge of capturing all factors influencing the treatment of experiments. Future research could address these limitations and further explore the complexities of news selection and reporting biases.

**Future Works.** Our methodology can be utilized and expanded in future research to quantify the spread of content from alternative social media platforms, such as Facebook and Instagram to news media. Additionally, a broader investigation could be conducted to measure the overall dissemination of content without limiting it to specific social movements. Moreover, by analyzing the Twitter data of news organizations, we can further explore and examine the tweets that are being retweeted and liked by these news outlets. This analysis will provide insights into how news organizations utilize their Twitter accounts to amplify tweets from other users.

Overall, our research provides valuable insights into the coverage of social movement tweets in news media, the factors influencing tweet selection, author selection bias, trends in reporting social movements, objectivity in reporting, and the presence of negative news coverage. These findings contribute to the understanding of how social movement-related content is disseminated and framed by news outlets, offering opportunities for further investigation and enhancing media transparency and accountability.

# Bibliography

[1] H. Abdi and L. J. Williams. Tukey's honestly significant difference (hsd) test. *Encyclopedia of research design*, 3(1):1–5, 2010.

[2] N. Alvarez-Gonzalez, A. Kaltenbrunner, and V. Gómez. Uncovering the limits of text-based emotion detection. In *Findings of the Association for Computational Linguistics: EMNLP*. Association for Computational Linguistics, 2021.

[3] J. An and H. Kwak. What gets media attention and how media attention evolves over time: Large-scale empirical evidence from 196 countries. In *Proc. Int. Conf. on Web and Social Media, ICWSM*. AAAI Press, 2017.

[4] M. Arango-Kure, M. Garz, and A. Rott. Bad news sells: The demand for news magazines and the tone of their covers. *Journal of Media Economics*, 27(4):199–214, 2014.

[5] N. Aruguete, E. Calvo, and T. Ventura. News sharing, gatekeeping, and polarization: A study of the# bolsonaro election. *Digital journalism*, 9(1):1–23, 2021.

[6] K. C. Bane. Tweeting the agenda: How print and alternative web-only news organizations use twitter as a source. *Journalism Practice*, 13(2):191–205, 2019.

[7] V. Basile, C. Bosco, E. Fersini, D. Nozza, V. Patti, F. M. R. Pardo, P. Rosso, and M. Sanguinetti. Semeval-2019 task 5: Multilingual detection of hate speech against immigrants and women in twitter. In *Proc. Int. Workshop on Semantic Evaluation, SemEval@NAACL-HLT*, pages 54–63. Association for Computational Linguistics, 2019.

[8] V. Belair-Gagnon, R. Zamith, and A. E. Holton. Role orientations and audience metrics in newsrooms: An examination of journalistic perceptions and their drivers. *Digital Journalism*, 8(3):347–366, 2020.

[9] D. Bernhardt, S. Krasa, and M. Polborn. Political polarization and the electoral effects of media bias. *Journal of Public Economics*, 92(5-6):1092–1104, 2008.

[10] N. Blanchett Neheli. News by numbers: The evolution of analytics in journalism. *Digital Journalism*, 6(8):1041–1051, 2018.

[11] B. J. Brands, T. Graham, and M. Broersma. *Social media sourcing practices: How Dutch newspapers use tweets in political news coverage*. Springer, 2018.

[12] W. Breed. Social control in the newsroom: A functional analysis. *Social forces*, pages 326–335, 1955.

[13] P. Bro and F. Wallberg. Digital gatekeeping: News media versus social media. *Digital journalism*, 2(3):446–454, 2014.

[14] M. Broersma and T. Graham. Twitter as a news source: How dutch and british newspapers used tweets in their news coverage, 2007–2011. *Journalism practice*, 7(4):446–464, 2013.

[15] J. Camacho-Collados, K. Rezaee, T. Riahi, A. Ushio, D. Loureiro, D. Antypas, J. Boisson, L. Espinosa-Anke, F. Liu, E. Martínez-Cámara, et al. TweetNLP: Cutting-Edge Natural Language Processing for Social Media. In *Proc. Conf. on Empirical Methods in Natural Language Processing: System Demonstrations*. Association for Computational Linguistics, 2022.

[16] J. J. Castle, S. Jenkins, C. D. Ortbals, L. Poloni-Staudinger, and J. C. Strachan. The effect of the# metoo movement on political engagement and ambition in 2018. *Political Research Quarterly*, 73(4):926–941, 2020.

[17] A. Chatterjee, K. N. Narahari, M. Joshi, and P. Agrawal. Semeval-2019 task 3: Emocontext contextual emotion detection in text. In *Proc. Int. Workshop on Semantic Evaluation*. Association for Computational Linguistics, 2019.

[18] K. Cho, B. van Merrienboer, Ç. Gülçehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In *Proc. Conf. on Empirical Methods in Natural Language Processing, EMNLP A meeting of SIGDAT, a Special Interest Group of the ACL*. ACL, 2014.

[19] J. W. Dearing and E. M. Rogers. *Agenda-setting*, volume 6. Sage, 1996.

[20] D. Demszky, D. Movshovitz-Attias, J. Ko, A. S. Cowen, G. Nemade, and S. Ravi. Goemotions: A dataset of fine-grained emotions. In *Proc. Annual Meeting of the Association for Computational Linguistics, ACL*. Association for Computational Linguistics, 2020.

[21] J. Devlin, M. Chang, K. Lee, and K. Toutanova. BERT: pre-training of deep bidirectional transformers for language understanding. In *Proc. Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019*, pages 4171–4186. Association for Computational Linguistics.

[22] C. L. Edrington. Social movements and identification: An examination of how black lives matter and march for our lives use identification strategies on twitter to build relationships. *Journalism & Mass Communication Quarterly*, 99(3):643–659, 2022.

[23] C. Emeran. The march for our lives movement in the united states. *When Students Protest: Secondary and High Schools*, page 145, 2021.

[24] G. Enli. Twitter as arena for the authentic outsider: Exploring the social media campaigns of trump and clinton in the 2016 us presidential election. *European journal of communication*, 32(1):50–61, 2017.

[25] A. K. Gautam, P. Mathur, R. Gosangi, D. Mahata, R. Sawhney, and R. R. Shah. #metooma: Multi-aspect annotations of tweets related to the metoo movement. In *Proc. Int. AAAI Conf. on Web and Social Media, ICWSM*. AAAI Press, 2020.

[26] A. Graves and J. Schmidhuber. Framewise phoneme classification with bidirectional lstm and other neural network architectures. *Neural networks*, 18(5-6):602–610, 2005.

[27] F. Hamborg and K. Donnay. Newsmtsc: A dataset for (multi-)target-dependent sentiment classification in political news articles. In *Proc. Conf. of the European Chapter of the Association for Computational Linguistics: Main Volume, EACL*. Association for Computational Linguistics, 2021.

[28] F. Hamborg, K. Donnay, and B. Gipp. Automated identification of media bias in news articles: an interdisciplinary literature review. *Int. J. Digit. Libr.*, 20(4):391–415, 2019.

[29] F. Hanusch. Web analytics and the functional differentiation of journalism cultures: individual, organizational and platform-specific influences on newswork. *Information, Communication & Society*, 20(10):1571–1586, 2017.

[30] F. Hanusch and E. C. Tandoc Jr. Comments, analytics, and social media: The impact of audience feedback on journalists' market orientation. *Journalism*, 20(6):695–713, 2019.

[31] T. Harcup and D. O'neill. What is news? news values revisited (again). *Journalism studies*, 18(12):1470–1488, 2017.

[32] J. Hartmann. Emotion english distilroberta-base. `https://huggingface.co/j-hartmann/emotion-english-distilroberta-base/`, 2022.

[33] N. Hassan, A. Poudel, J. G. Hale, C. Hubacek, K. T. Huq, S. K. K. Santu, and S. I. Ahmed. Towards automated sexual violence report tracking. In *Proc. Int. AAAI Conf. on Web and Social Media, ICWSM*. AAAI Press, 2020.

[34] M. Heinz and J. F. Swinnen. Media bias in economic news: A factor 20. *Available at SSRN 2281103*, 2013.

[35] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Comput.*, 9(8):1735–1780, 1997.

[36] N. Houlsby, A. Giurgiu, S. Jastrzebski, B. Morrone, Q. de Laroussilhe, A. Gesmundo, M. Attariyan, and S. Gelly. Parameter-efficient transfer learning for NLP. In *Proc. Int. Conf. on Machine Learning, ICML*. PMLR, 2019.

[37] M. Hu and B. Liu. Mining and summarizing customer reviews. In *Proc. ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*. ACM, 2004.

[38] T. Hu, M. H. Ribeiro, R. West, and A. Spitz. Quotatives indicate decline in objectivity in U.S. political news. *CoRR*, abs/2210.15476, 2022.

[39] M. Jeon, H. O. Kim, and C. W. Woo. # metoo movement in political media era: a comparison of us media and korean media. *Communication Quarterly*, 70(1):22–41, 2022.

[40] J. Külz, A. Spitz, A. Abu-Akel, S. Günnemann, and R. West. United states politicians' tone became more negative with 2016 primary campaigns. *CoRR*, abs/2207.08112, 2022.

[41] D. L. Lasorsa, S. C. Lewis, and A. E. Holton. Normalizing twitter: Journalism practice in an emerging communication space. *Journalism studies*, 13(1):19–36, 2012.

[42] C. S. Lee and L. Ma. News sharing in social media: The effect of gratifications and prior experience. *Comput. Hum. Behav.*, 28(2):331–339, 2012.

[43] X. Li and P. Li. Rejection sampling for weighted jaccard similarity revisited. In *Thirty-Fifth AAAI Conf. on Artificial Intelligence, AAAI , The Eleventh Symp. on Educational Advances in Artificial Intelligence, EAAI*. AAAI Press, 2021.

[44] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. Roberta: A robustly optimized BERT pretraining approach. *CoRR*, abs/1907.11692, 2019.

[45] M. Luca, D. Malhotra, and C. Poliquin. The impact of mass shootings on gun policy. *Journal of public economics*, 181:104083, 2020.

[46] S. Mohammad, F. Bravo-Marquez, M. Salameh, and S. Kiritchenko. Semeval-2018 task 1: Affect in tweets. In *Proc. Int. workshop on semantic evaluation*, 2018.

[47] S. Mohammad and P. Turney. Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In *Proc. of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text*, 2010.

[48] S. J. Moon and P. Hadley. Routinizing a new technology in the newsroom: Twitter as a news source in mainstream media. *Journal of Broadcasting & Electronic Media*, 58(2):289–305, 2014.

[49] I. Oelrichs. Just copy and paste? usage and patterns of social media sources in online articles on sport. *International Journal of Sport Communication*, 15(4):325–335, 2022.

[50] C. Oschatz, S. Stier, and J. Maier. Twitter in the news: An analysis of embedded tweets in political news coverage. *Digital Journalism*, 10(9):1526–1545, 2022.

[51] S. Paulussen and R. A. Harder. Social media references in newspapers: Facebook, twitter and youtube as sources in newspaper journalism. *Journalism practice*, 8(5):542–551, 2014.

[52] J. W. Pennebaker, M. E. Francis, and R. J. Booth. Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001):2001, 2001.

[53] J. Pfeiffer, A. Rücklé, C. Poth, A. Kamath, I. Vulic, S. Ruder, K. Cho, and I. Gurevych. Adapterhub: A framework for adapting transformers. In *Proc. Conf. on Empirical Methods in Natural Language Processing: System Demonstrations, EMNLP.* Association for Computational Linguistics, 2020.

[54] C. Poth, J. Pfeiffer, A. R"uckl'e, and I. Gurevych. What to pre-train on? Efficient intermediate task selection. In *Proc. Conf. on Empirical Methods in Natural Language Processing.* Association for Computational Linguistics, 2021.

[55] N. Reimers and I. Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proc. Conf. on Empirical Methods in Natural Language Processing and the 9th Int. Joint Conf. on Natural Language Processing, EMNLP-IJCNLP.* Association for Computational Linguistics, 2019.

[56] P. R. Rosenbaum, P. Rosenbaum, and Briskman. *Design of observational studies*, volume 10. Springer, 2010.

[57] S. Rosenthal, N. Farra, and P. Nakov. Semeval-2017 task 4: Sentiment analysis in twitter. *CoRR*, abs/1912.00741, 2019.

[58] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986.

[59] M. J. Salganik. *Bit by bit: Social research in the digital age*. Princeton University Press, 2019.

[60] M. Schudson. The sociology of news production. *Media, culture & society*, 11(3):263–282, 1989.

[61] H. A. Semetko and K. Schoenbach. News and elections: German bundestag campaigns in the bild, 1990-2002. *Harvard International Journal of Press/Politics*, 8(3):54–69, 2003.

[62] P. J. Shoemaker and T. Vos. *Gatekeeping theory*. Routledge, 2009.

[63] J. B. Singer. On a role: online newspapers, participatory journalism, and the us presidential elections. In *Gatekeeping in transition*, pages 85–103. Routledge, 2015.

[64] L. Tan, S. Ponnam, P. Gillham, B. Edwards, and E. Johnson. Analyzing the impact of social media on social movements: a computational study on twitter and the occupy wall street movement. In *Advances in Social Networks Analysis and Mining, ASONAM*. ACM, 2013.

[65] E. C. Tandoc Jr. Journalism is twerking? how web analytics is changing the process of gatekeeping. *New media & society*, 16(4):559–575, 2014.

[66] E. C. Tandoc Jr. Journalism is twerking? how web analytics is changing the process of gatekeeping. *New media & society*, 16(4):559–575, 2014.

[67] E. C. Tandoc Jr and T. P. Vos. The journalist is marketing the news: Social media in the gatekeeping process. *Journalism practice*, 10(8):950–966, 2016.

[68] M. Trussler and S. Soroka. Consumer demand for cynical and negative news frames. *The International Journal of Press/Politics*, 19(3):360–379, 2014.

[69] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems*, 2017.

[70] T. Vaucher, A. Spitz, M. Catasta, and R. West. Quotebank: A corpus of quotations from a decade of news. In *WSDM '21, Int. Conf. on Web Search and Data Mining*. ACM, 2021.

[71] G. Von Nordheim, K. Boczek, and L. Koppers. Sourcing the sources: An analysis of the use of twitter and facebook as a journalistic source over 10 years in the new york times, the guardian, and süddeutsche zeitung. *Digital journalism*, 6(7):807–828, 2018.

[72] D. Vrandecic and M. Krötzsch. Wikidata: a free collaborative knowledgebase. *Commun. ACM*, 57, 2014.

[73] H. T. Vu. The online audience as gatekeeper: The influence of reader metrics on news editorial selection. *Journalism*, 15(8):1094–1110, 2014.

[74] K. Welbers, W. Van Atteveldt, J. Kleinnijenhuis, N. Ruigrok, and J. Schaper. News selection criteria in the digital age: Professional norms versus online audience metrics. *Journalism*, 17(8):1037–1053, 2016.

[75] P. J. Werbos. Backpropagation through time: what it does and how to do it. *Proc. IEEE*, 78(10):1550–1560, 1990.

[76] T. Wilson, J. Wiebe, and P. Hoffmann. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proc. of human language technology conf. and conf. on empirical methods in natural language processing*, pages 347–354, 2005.

[77] R. F. Woolson. Wilcoxon signed-rank test. *Wiley encyclopedia of clinical trials*, pages 1–3, 2007.

[78] T. Zaman, E. B. Fox, and E. T. Bradlow. A bayesian approach for predicting the popularity of tweets. 2014.

[79] M. Zampieri, S. Malmasi, P. Nakov, S. Rosenthal, N. Farra, and R. Kumar. Semeval-2019 task 6: Identifying and categorizing offensive language in social media (offenseval). In *Proc. Int. Workshop on Semantic Evaluation, SemEval@NAACL-HLT*. Association for Computational Linguistics, 2019.

[80] H. M. Zoller and D. Casteel. # march for our lives: health activism, diagnostic framing, gun control, and the gun industry. *Health communication*, 37(7):813–823, 2022.

CHAPTER 7

# Appendix

## 7.1 Appendix A: Performance of Classifiers

The performance of the classification models in predicting whether a tweet will be reported by news outlets is displayed in Table 15. Given the scarcity of reported tweets, a random oversampling technique was used on the training dataset.

**Table 15** Accuracy and F1-score of machine learning models trained to predict whether a tweet gets reported.

| Social Movement -User group | Classifier | Accuracy | F1-Score |
|---|---|---|---|
| MeToo-Celebrity | SVM | 0.93 | 0.61 |
| | Random Forest | 0.96 | 0.58 |
| | Gradient Boost | 0.91 | 0.58 |
| | Ada Boost | 0.91 | 0.56 |
| MeToo-Politician | SVM | 0.96 | 0.53 |
| | Random Forest | 0.98 | 0.56 |
| | Gradient Boost | 0.95 | 0.62 |
| | Ada Boost | 0.93 | 0.68 |
| MFOL-Celebrity | SVM | 0.96 | 0.53 |
| | Random Forest | 0.98 | 0.54 |
| | Gradient Boost | 0.96 | 0.55 |
| | Ada Boost | 0.96 | 0.57 |
| MFOL-Politician | SVM | 0.98 | 0.53 |
| | Random Forest | 0.99 | 0.49 |
| | Gradient Boost | 0.96 | 0.56 |
| | Ada Boost | 0.97 | 0.57 |