Classifying News Articles Following the 2020 Presidential Election by Tone of Language and Media Bias

Rachel Lee, Angie Yang

December 16, 2020

1 Introduction

The intended purpose of news is to publish and distribute a factual account of recent events. News outlets hold immense power in shaping the global political narrative, thus impacting audience's understandings of reality. In the wake of controversial politics and increasing misinformation, there is not one singular political narrative. The growing disparity between popular political narratives makes it difficult for any source of information to be entirely neutral, especially when there is dispute on facts in and of itself. From an author's personal biases to a news organization's political leanings, it is likely for news articles to convey these sentiments. Whether it be bias from word choice and labeling (WCL bias [3]) or an extreme interpretation of reality, the majority of news reports are unlikely be free of any political leaning. Many will claim that even the most reputable news outlets have taken a political stance, which then bleeds into their news coverage. Indeed, we as authors of this scientific paper will also face this challenge as we write on this issue.

Just as conservatives and liberals might have different attitudes and ways of speaking about certain topics, right wing and left wing news outlets also frame political concepts in their own terms. Framing is an inescapable aspect of communication which is defined as selecting "a few elements of perceived reality and assembling a narrative that highlights connections among them to promote a particular interpretation" [2]. For example, take the controversial topic of abortion. Conservative news sources would call their stance "pro-life" while liberal news sources would label their belief as "pro-choice". As one can see, these words are both positive. As each side justifies their stance through persuasive language, the reader begins to see the morality in the news they are reading. Conservative papers might describe abortion as "killing an innocent baby", whereas liberal news might describe abortion as "getting rid of a clump of cells by the woman's choice". One can imagine how different articles from two news outlets can be justify their own stance. Other examples of this include using the phrase "global warming" as opposed to "climate change" [10] and also "gay civil unions" vs. "homosexual marriage" [9].

Another factor that comes into play is confirmation bias. People are most likely to read news articles that support their existing views. Most of the time, people do not want to be challenged and just want to read news that follow what they already believe. Thus, to appeal to their intended audiences, news articles may decide to only reiterate and strengthen what people are already thinking. The language of news articles have greatly influence how individuals interpret current events. How news organizations describe the future president can have great implications on how the president will be received by Americans and possibly populations outside of America as well.

Through this project, we aim to assess the existing biases of news outlets using the specific example of news reports published after the 2020 Presidential Election results. As expected, throughout the campaigning period, many liberal leaning news outlets painted Joe Biden in a positive light and many conservative leaning news outlets spoke negatively about Joe Biden. We would like to confirm or perhaps challenge our hypothesis that this trend will continue from the campaigning period to after the results were published. Since Joe Biden was ultimately elected as the next president [7], we expect news sources typically considered conservative to produce news articles on the election results with a negative tone, and news sources typically considered liberal to produce election results news articles with a positive tone.

2 Related Works

There are a multitude of ways to approach this problem. Researchers have used recursive neural networks on various Congressional debates to detect political ideology [6]. In this study, Iyyer et al. discovered that recursive neural networks were more effective in classifying political ideology as opposed to the simpler bag-of-words models and also hand-designed lexica. In addition to that, they also analyzed the ideas that each political party brought up the most, conservatives emphasizing freedom and religion, and liberals focusing on the gap between the rich and the poor, minorities, and the working class.

Another study focused on automating the process of identifying the frame of a news coverage through WCL bias [4]. Hamborg employs cross-document coreference resolution (CDCR) and target-dependent sentiment classification (TSC) including "sentiment shift" and identification of framing effects and causes to see if machine learning processes can detect frame as well as the established manual process. Using 3 different news datasets, the study was able to correctly identify frame with an average recall of 70%. Hamborg concludes that an automated method of labeling WCL bias is useful and effective to implement, noting that integrating a media bias score into r news aggregators would be beneficial for informing audiences.

Lin et al., conducted a study using news articles regarding the Israeli-Palestinian conflict to create a classification model to determine the bias of the writer [8]. First, the researchers built the model to classify each document at a time. They constructed multiple different models such as Naive Bayes and Support Vector Machines (SVM) with a goal to compare models like Naive Bayes that generate and models like Support Vector Machines that discriminate. They found the Naive Bayes, reaching an accuracy rate of 99%, yielded higher accuracy rates than SVM, with a 97% accuracy when classifying documents.

The researchers then went on to examine deeper to each sentence. Using a Latent Sentence Perspective Model (LSPM), they compared this model with the Naive Bayes model, classifying each sentence as a time. The LSPM with an accuracy of 95% slightly outperformed the Naive Bayes model which had an accuracy of 93%. Ultimately, they discovered that the accuracy of the classification models when classifying sentences (94%) was significantly lower than the accuracy of classifying documents (99%). This is understandable, as the machine has more information to base its decision on, making its decision more accurate when there is more data, in an entire document.

In a study conducted by Hardisty et al., researchers trained and tested their model on the same data set as Lin et al., determining bias based on documents about the Israeli-Palestinian conflict [5]. They predicted that strings of words longer than n-grams could potentially be useful in determining political bias in a text. To test their theory, they constructed a non-parametric Bayesian model which has more flexibility to change in size according to the data. This adapted Naive Bayes classifier used adaptor grammars that simplify expressing non-parametric Bayesian models through context free rules in letting nonterminals, or symbols that can be erased, be rewritten to subtrees.

The difference that Hardisty et al. discovered between using a full Naive Bayes model (93%) and using the adapted Naive Bayes model (99%) was around 6%. Since the non-parametric Naive Bays model adapted model to the data, there was not a one size fits all system that worked best for every data set. This way, the program is able to cater to the data that is is classifying and maximize accuracy in this way.

3 Data Collection

Data is collected from 6 different news sources: Newsweek, MSNBC, NPR News, USA Today, Breitbart, and The Spectator USA (also known as The American Spectator). These specific news sources are selected to be representative of the range media biases that exist. Based on the AllSides Media Bias Rating published in 2020 [1] , these news sources were rated as follows:

Newsweek and MSNBC are political left; NPR News and USA Today are political center; and Breitbart and The American Spectator are political right. By using news articles from these varying news organizations, we aim to produce a data set that would be fairly representative of the political bias spectrum in media.

For each of the above news sources, we search google news with the query term of the news organization's name followed by ”Election Results”. This is performed on a private browsing session so search results are not be skewed by our previous browser activity. We narrow the time period to be between November 7, 2020, the date on which the Associated Press announced Joe Biden as the president elect [7] , and November 14, 2020. The first 10 articles that appear are the articles that we use for our data set. The text from these articles are saved in files to be used for our algorithm.

The text is also fed into Grammarly. Utilizing the beta tone detection feature that Grammerly has recently introduced, we were able to characterize the tone of language each article used to be either positive (ex. optimistic, admiring), negative (ex. worried, sad), or neutral (ex. formal, informative). These labels are used for our algorithm to classify tone based on word choice. We also replicate this data set but with labels of media bias according to AllSides [1] . This is used for our algorithm to classify media bias based on word choice.

4 Approach and Model Description

The tone of the article as well as the word choice both result from the political bias of a particular journalist or news outlet. Many times, one is able to differentiate the political standing of a certain news outlet by only reading an article. While creating a model that classifies news articles into different categories, we first train on news articles categorized by the tone as classified by Grammerly. After this, we use our same model to classify the news articles based on the political standing of the news outlet, as determined by AllSides. Our goal in attempting to classify articles into media bias is to see whether word choice could be useful in directly determining the media bias of a certain article after testing word choice in determining tone.

In our model that we use for both classifying tone as well as classifying political bias, we first pre-process the words in the article to be useful in the machine learning process (converting all words to lowercase and filtering out stopwords). When all the words are ready to process, we label each article with the categories we were trying to classify, so the Grammerly tone in classifying tone, and the political bias in classifying media bias. After this, we randomly shuffle the articles and choose the 2000 most frequent unigrams and bigrams in all the news articles as features in our model. Finally, we use a Naive Bayes Classifier to train the model with 10/60 articles (16.7%) and test the accuracy of the model in classifying the news articles with the remaining 50 articles.

5 Results

We conduct 30 trials of our algorithm with each of the 2 data set labelings. With articles labeled with the Grammerly detected tone, the average accuracy is 58.7% and the standard deviation is 13.6. With articles labeled with AllSides' media bias rating, the average accuracy is 62.7% is the standard deviation is 16.1. Below is a graph of the accuracy scores (Fig 3).

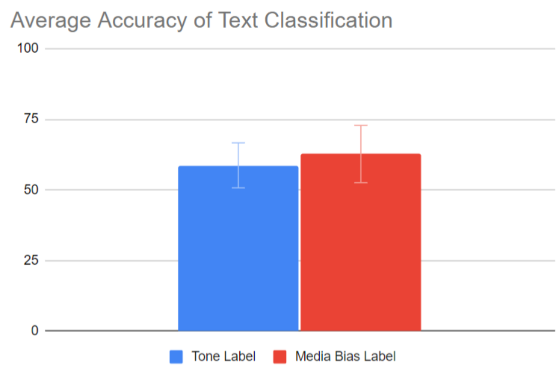

Figure 1: For both tone labeling and media bias labeling, we averaged the accuracy scores over 30 trials. Error bars represent 1 standard deviation.

For each of the trials, we also output the 15 most informative features among both unigrams and bigrams. Below tables compiling the 10 features that appeared the most for trials where we classify by tone labels (Fig 2) and where we classify by media bias labels (Fig 3). Note that we specifically excluded stop words from being features, since they are so common in language and we can assume that they do not have a political bias.

Word	Feature Details	Ratio	Frequency in Top 15 Informative Features
Everyone	contains(everyone) = True	neg : pos = 9.9 : 1	0.80
Took	contains(took) = True	neg : pos = 8.9 : 1	0.53
Period	contains(period) = True	neg : pos = 8.6 : 1	0.63
Agreement	contains(agreement) = True	neg : pos = 8.4 : 1	0.67
Looked	contains(looked) = True	neg : pos = 8.3 : 1	0.67
Simply	contains(simply) = True	neg : pos = 8.1 : 1	0.73
Reform	contains(reform) = True	neg : pos = 8.1 : 1	0.57
Blue	contains(blue) = True	neg : pos = 8.1 : 1	0.60
Pandemic	contains(pandemic) = True	neg : pos = 7.9 : 1	0.37
Accusations	contains(accusations) = True	neg : pos = 7.7 : 1	0.57

Figure 2: Most Informative Features for Classifying by Tone Labels

Word	Feature Details	Ratio	Frequency in Top 15 Informative Features
Vice	contains(vice) = True	center : left = 6.1 : 1	0.97
What	contains(what) = True	center : right = 5.8 : 1	0.43
Federal	contains(federal) = True	center : right = 5.5 : 1	0.53
Win	contains(win) = False	right : center = 5.2 : 1	0.70
After	contains(after) = True	left : center = 5.0 : 1	0.47
Incumbent	contains(incumbent) = True	left : right = 4.9 : 1	0.37
Back	contains(back) = True	left : right = 4.9 : 1	0.90
Director	contains(director) = True	center : right = 4.8 : 1	0.43
Saying	contains(saying) = True	center : left = 4.7 : 1	0.47
Continues	contains(continues) = True	center : right = 4.2 : 1	0.33

Figure 3: Most Informative Features for Classifying by Media Bias Labels

6 Analysis

Our results for both classification by tone labeling and media bias labeling are not very promising. The accuracy scores are low and the standard deviations are high, meaning that we have both low accuracy and low precision. One positive sign is that our average accuracy scores are above 33%, indicating that the trained classifier is performing at better than random chance, since we have 3 classes in each trial. The accuracy score for classifying media bias labels is higher than classifying tone labels. This could indicate that our features are better suited to identifying media bias, and that media bias is more directly revealed in word choice than tone is. However, we likely require further experimentation and more accurate results in order to support that claim.

Other experimentation that we would like to explore, if given the time, would be other classification methods. Our current algorithm uses a naive bayes classifier. In our research, we read about similar studies using recursive neural networks (RNN) and Support Vector Machines (SVM). Using these different classification techniques on the data set we have might produce better results as they could be better suited to the text classification in this situation.

In addition to accuracy results, the informative features are interesting, specifically because they only include unigrams. Within the top 15 informative features of 30 trials for both tone and media bias, we never once see a bigram. This indicates that individual words have a stronger relationship to the classification of tone and media bias of a text than pairs of words. In further experimentation, we might try to alter the feature extraction to produce more informative features. The top informative features also includes many words that do not seem like they would have a bias. For example in our tone classification trials, ”everyone” was the most frequent word that appeared in the top 15 informative features. However, this could also be attributed to a small sample size.

A serious shortcoming of our experimental results is our limited sample size. Our data set only included 60 articles in total, 20 from each media bias class. Class imbalance is also a serious issue for the tone classes as we had an overwhelming number of positive tone articles and very few neutral and negative. There were 38 articles with positive tone and only about 10 for neutral and negative out of the 60 total articles. (A possible reason for this overwhelming positive tone across the media bias spectrum will be discussed in the conclusion.)

We can see indicators of this class imbalance being a problem when we look at the most informative features (Fig 2). All of the ratios are between negative and positive. If compared with the media bias experimental results that did not have this class imbalance issue, it is easy to see that among the most informative features, the ratios contain a mix of different class combinations (Fig 3). This class imbalance can be addressed through increasing our sample size. We were not able to do this within the given time frame, as data collection was all done manually, making it extremely time consuming. Another potential solution could be re-sampling our existing data by replicating instances of the under-represented classes (neutral and negative tone) or deleting instances of over-represented classes (positive tone). However this would likely not be as effective as expanding our data set.

Overall, our results revealed that word choice did not accurately predict media bias, however we note that more experimentation is required. Our current algorithm using Naive Bayes classification on data sets labeled with tone and media bias classes produced low accuracy and low precision. Improvements that could be explored are trying different classification methods, as well as expanding our data set.

7 Conclusion

From collecting news articles, categorizing them by tone and media bias to writing our own code and building our own model from scratch, we conclude that tone in language may not be an effective way to determine media bias, especially in the context of news articles in the time frame immediately following the news of President-elect Joe Biden claiming the race over the current President, Donald Trump.

While collecting data and detecting tone from each news article, we find that many right-wing news outlets have very positive and optimistic sounding articles, contrary to what we had previously predicted. Many of these articles were still optimistic that Trump would win the election even after the Associated Press had declared Biden the winner. It seems like there are two completely opposite political narratives that are prominent in America; however, both are optimistic towards a future that fits their own ideals. This use of tone was not accounted for in our predictions that tone would indicate media bias.

News outlets have a large amount of influence on how their readers view certain issues and what they believe in. Of course, readers often choose to support and read news outlets that align with their political standing, but reading these articles also perpetuate and strengthen these existing beliefs. As the country grows increasingly divided between liberals and conservatives, it is reflected in the popular news media that is consumed by a significant portion of Americans. Although our nation has become more divided, from the news articles we collected, perhaps Americans are also becoming more hopeful, looking forward to an ideal future as they believe.

We imagine that this played a major role in the results we achieved with our classifier. Since articles from both ends of the political spectrum had similar tones and sentiments, it is especially difficult to construct an accurate model that successfully differentiates the political biases of the article.

If we had more time we would collect more data to provide the model with more information so that it can possibly create a more informed and accurate decision. We would also try more complex classification models such as a recursive neural network with varying depths and breadths to see if it would yield a more effective classifier, and we could draw more concrete conclusions.

8 Group Work Breakdown

1. Introduction: Angie and Rachel
2. Related Works: Angie and Rachel
3. Data Collection: Rachel
4. Approach and Model Description: Angie
5. Results: Rachel
6. Analysis: Rachel
7. Conclusion: Angie
8. Coding: Angie and Rachel

References

[1] Media bias ratings, 2020.

[2] Robert M. Entman. Framing bias: Media in the distribution of power. *Journal of Communication*, 57(1):163–173, 2007.

[3] F. Hamborg, A. Zhukova, and B. Gipp. Automated identification of media bias by word choice and labeling in news articles. In *2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, pages 196–205, 2019.

[4] Felix Hamborg. Media bias, the social sciences, and NLP: Automating frame analyses to identify bias by word choice and labeling. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*, pages 79–87, Online, July 2020. Association for Computational Linguistics.

[5] Eric Hardisty, Jordan Boyd-Graber, and Philip Resnik. Modeling perspective using Adaptor Grammars. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 284–292, Cambridge, MA, October 2010. Association for Computational Linguistics.

[6] Mohit Iyyer, Peter Enns, Jordan Boyd-Graber, and Philip Resnik. Political ideology detection using recursive neural networks. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1113–1122, Baltimore, Maryland, June 2014. Association for Computational Linguistics.

[7] Zeke Miller Jonathan Lemire and Will Weissert. Biden defeats trump for white house, says 'time to heal'. 2020.

[8] Wei-Hao Lin, Theresa Wilson, Janyce Wiebe, and Alexander Hauptmann. Which side are you on? identifying perspectives at the document and sentence levels. In *Proceedings of the Tenth Conference on Computational Natural Language Learning (CoNLL-X)*, pages 109–116, New York City, June 2006. Association for Computational Linguistics.

[9] Vincent Price, Lilach Nir, and Joseph N. Cappella. Framing public discussion of gay civil unions. *The Public Opinion Quarterly*, 69(2):179–212, 2005.

[10] Jonathon Schuldt, Sara Konrath, and Norbert Schwarz. Global warming' or 'climate change'?: Whether the planet is warming depends on question wording. *The Public Opinion Quarterly*, 75:115–124, 03 2011.