

# Detecting Frames in News Headlines and Its Application to Analyzing News Framing Trends Surrounding U.S. Gun Violence

Siyi Liu   Lei Guo   Kate Mays   Margrit Betke   Derry Tanti Wijaya

Boston University

{liusiyi, guolei, kkmays, betke, wijaya}@bu.edu

## Abstract

Different news articles about the same topic often offer a variety of perspectives: an article written about gun violence might emphasize gun control, while another might promote 2nd Amendment rights, and yet a third might focus on mental health issues. In communication research, these different perspectives are known as “frames”, which, when used in news media will influence the opinion of their readers in multiple ways. In this paper, we present a method for effectively detecting frames in news headlines. Our training and performance evaluation is based on a new dataset of news headlines related to the issue of gun violence in the United States. This Gun Violence Frame Corpus (GVFC) was curated and annotated by journalism and communication experts. Our proposed approach sets a new state-of-the-art performance for multiclass news frame detection, significantly outperforming a recent baseline by 35.9% absolute difference in accuracy. We apply our frame detection approach in a large scale study of 88k news headlines about the coverage of gun violence in the U.S. between 2016 and 2018.

## 1 Introduction

The political climate in the United States is increasingly polarized (Pew Research Center, 2018a). To many media scholars and pundits, the main reason that liberals and conservatives inhabit different worlds is that news media of varied political orientations have been depicting two distinct versions of social reality (Mitchell et al., 2014; Stroud, 2011). To address this problem, one needs to assess the ways in which news reporters frame important public affairs. In communication research, “to frame” means “to select some aspects of a perceived reality and make them more salient in a communicating text” (Entman, 1993). Like any type of communication, news involves

framing. In a polarized media environment, partisan media outlets intentionally frame news stories in a way to advance certain political agendas (Jamieson et al., 2007; Levendusky, 2013). Even when journalists make their best efforts to pursue objectivity, media framing often favors one side over another in political disputes, thus always resulting in some degree of bias (Entman, 2010). Hence, a news framing analysis is helpful because it not only tells us whether a news article is left- or right-leaning (or positive or negative), but also reveals how the article is structured to promote a certain side of the political spectrum.

In communication research, manual identification of media frames is a challenging task due to the large amount of media data in this news-saturated environment. More importantly, there is a high level of complexity in framing analysis that often requires a careful investigation of nuances in news coverage, which is time-consuming. In the field of Natural Language Processing (NLP), automated news framing analysis is a relatively unexplored area. Existing sentiment-analysis techniques fall short of addressing the nuances needed for framing analysis, which requires the detection of perspectives beyond positive and negative.

In this paper, we develop a neural network based approach for classifying frames in news article headlines by fine-tuning a state-of-the-art language representation model (BERT: Bidirectional Encoder Representations from Transformers (Devlin et al., 2018)) for the task of frame detection.

Here, we focus on the application of news frame detection on one prominent public affairs issue in the United States, namely, gun violence. Some of the deadliest mass shootings have happened during the past few years. In fact, the United States has the highest rate of gun-related homicides in the developed world. However, Republicans and Democrats remain divided on whether gun vio-

lence is an important issue and disagree on most gun-related policies, making gun violence one of the most polarized issues in the country. (Pew Research Center, 2018b). As a result, despite the seriousness of the issue in reality, it is not considered a priority that should be tackled at the Congressional level. One factor that potentially explains the divergence of public opinion is how different politically oriented news media cover gun violence. It is likely that liberal-leaning and conservative-leaning media frame the issue in different ways, which may ultimately determine different publics' perception of the issue.

We use our frame detection approach to automatically detect frames of news article headlines related to gun violence during the past few years, which enables large scale analysis of framing trends surrounding this issue in the United States. Specifically, we focus on the years 2016, 2017, and 2018 because these three years have witnessed a number of high-profile mass shootings, which often reignited national gun debate.

Overall, our analysis results in interesting findings about U.S. media coverage of gun violence that speak to the divided media and political landscape in the country. Our contributions are twofold: Firstly, we have developed a state-of-the-art news frame detection approach by fine-tuning BERT language model to perform the multiclass (frame) classification on news article headlines. Our approach significantly outperforms a recent baseline in automated news frame detection (Field et al., 2018) and other neural network baselines.

Secondly, we have curated a new dataset of news articles related to U.S. gun violence: the Gun Violence Frame Corpus (GVFC), which contains news headlines and their frame annotations from 21 major U.S. news organizations. This dataset is the first of its kind in that it is carefully curated and contains domain-expert annotations of frames in news headlines. We use our model trained on GVFC to do a large scale analysis of U.S. gun violence framing trends in the U.S. between 2016 and 2018.

## 2 Related Work

### 2.1 News Framing

Framing is a subtle form of media manipulation in which some aspects of a topic are highlighted in order to promote a particular interpretation. It is related to the word choice and labeling by jour-

nalists (Hamborg et al., 2019) for example, by choosing “illegal alien” instead of “undocumented immigrant”, journalists can highlight different aspects of an immigration issue.

Communication researchers have developed a variety of approaches to analyzing media framing. One popular quantitative approach is to first identify a list of frames and then manually classify news articles into one of the identified frames. Journalists often use generic frames that are common across a range of issues, such as human interest, conflict, and economic consequences (Russell Neuman et al., 1993; Nisbet, 2010; Semetko and Valkenburg, 2000), on top of issue-specific frames in their reporting. There are a number of issue-specific frames that have been particularly related to the issue of gun violence in the United States. On a basic level, the debate about guns has been framed as a threat to *public safety* (Haider-Markel and Joslyn, 2001; Lawrence and Birkland, 2004), enabled by weak *gun laws* (Birkland and Lawrence, 2009), versus an individual right to have access to guns secured by the *2nd Amendment's* “right to bear arms” (Haider-Markel and Joslyn, 2001). Lawrence and Birkland (2004); Birkland and Lawrence (2009) also described how, after the Columbine shooting, the media discourse framed violent *popular culture* (e.g., movies and video games that glorify violence) as a culprit. Beyond the issue itself, the debate surrounding gun violence has also been framed as a Democrat vs. Republican *political contest* (Schnell, 2001).

In health communication, researchers have also examined the extent to which the news media frame the issue from the perspective of “dangerous people” (e.g., those with mental illness) wielding weapons as compared to “dangerous weapons” (e.g., large-capacity assault rifles) causing gun violence (McGinty et al., 2013). The *mental illness* of gunmen is often a focal point in the coverage of mass shootings (McGinty et al., 2014). Related to the issue of mental health are broader concerns about troubled individuals who lack the social support and resources to receive the help that they need (DeFoster and Swalve, 2018). The discussion about *race and ethnicity* has also emerged as a salient frame, in that news coverage of gun violence may differ somewhat depending on who the perpetrators are (Leavy and Maloney, 2009).

For our dataset, we detect these issue-specific frames typically found in media coverage of gun

violence, as well as generic frames like *economic consequences*. While it would be beneficial to automate the framing analysis across a variety of issues, here we argue that developing issue-specific tools will allow for a more nuanced understanding of each issue. Using our approach of combining expert-chosen frames for a particular issue and an automatic detection of these frames, communication researchers can further investigate how different news media influence public opinion in subtle ways *and* at scale, and thus be able to help prepare stronger arguments for journalistic practice and ultimately policy changes about the issue.

## 2.2 News Frame Detection

Media Frame Corpus (MFC) (Card et al., 2015) is one of the first large-scale datasets of frame annotations. It contains 11,900 hand-annotated English news articles for media framing that cover three issues: immigration, tobacco, and same-sex marriage. Undergraduate student annotators highlight the span of text that covers a frame following an annotation codebook. MFC has 15 generic media frames, which are defined in the *Policy Frames Codebook* by Boydston et al. (2014), such as economics, political, quality of life, and also an “other” label for news articles that cannot be covered by any of the 15 frames. These news articles have been collected using keyword search from 13 national U.S. newspapers from 1990 to 2012 and contains 38,283 news articles. Duplicate and near-duplicate articles were removed and 20,037 of these articles were randomly selected for manual framing annotation. Aside from spans of text, headlines and entire news text are also annotated with the headline and primary frames respectively.

Naderi and Hirst (2017) detect news frames at the sentence-level using deep recurrent neural networks, specifically LSTM, BiLSTM, and GRU. They used news articles from MFC dataset (Card et al., 2015) to train and evaluate their model. They show that their results for frame detection are better than classifiers that rely on topics models for detecting frames (Tsur et al., 2015; Nguyen et al., 2015). Our work is different from theirs in that we focus on detecting the frame in the news article headline, which unlike a complete sentence, is typically a short phrase. We implement these deep recurrent networks in our experiments as baselines and find that our approach performs better for detecting frames in headlines, both in MFC and our

GVFC. We also implement a recent word-based method for detecting frames in English and Russian news articles (Field et al., 2018) as another baseline. We detail these baseline approaches and their results in our experiment section (section 4).

## 3 Dataset Creation

### 3.1 News Article Collection

We drew our sample of news articles from a list of top U.S. news websites defined in terms of traffic to the websites. We cross-referenced several sources that had “top news sites” of their own: the Pew Research Center (2018b), Statista (2017), Alexa (2018), and MediaCloud, which is an open-source online platform. We synthesized these lists towards creating one list that contained news sites from the left, center, and right sides of the ideological spectrum based on categories defined in MediaCloud; Pew Research Center (2016); Ad Fontes Media (2019). We started with list of 30 media outlets based on these references.

We collected articles from these outlets from four time periods over the course of 2018 in order to capture a diversity of articles. Some articles were collected over periods during or immediately after a mass shooting (e.g., the Parkland School shooting in 02/2018). Other articles were collected when gun violence was not necessarily the most salient current event. We also included articles from several months before the 2018 U.S. midterm elections as the gun-related issue was a central topic for political discussion during this period. The articles were retrieved using Crimson Hexagon’s ForSight social media analytics platform (Hexagon, 2018), retrieving articles that had at least one keyword in their headlines from the following list: {“gun,” “firearm,” “NRA,” “2nd amendment,” “second amendment,” “AR15,” “assault weapon,” “rifle,” “Brady act,” “Brady bill,” “mass shooting”}. We came up with the list of keywords based on the previous literature and on the review of a sample of our data. After collecting the articles, news articles with duplicate titles were removed and the rest sampled to be analyzed and annotated. After sampling and annotation, the final dataset contains frame annotations of news articles from a total of 21 media outlets<sup>1</sup>.

<sup>1</sup>For reproducibility and future research, we make our dataset and annotation codebook publicly available at <https://derrywijaya.github.io/GVFC.html>

### 3.2 News Article Annotation

Quantitative content analysis (QCA) in communication research is a commonly used method to derive “replicable and meaningful inferences from texts (or other meaningful matter)” (Krippendorff, 2004). To perform QCA, one draws a representative sample of text (or other types of content), on which two or more trained *coders* (i.e., annotators) apply a codebook protocol, which should include all of the variables for annotation and their definitions. Prior to coding the entire sample independently, coders are first trained on the codebook, and their agreement on how to apply the codes is measured with inter-coder reliability (ICR). High ICR values implies that two or more coders consistently categorized the content similarly, which signals a high validity of the coded results. Once coders have reached an acceptable ICR (above 90% agreement or 0.70 Krippendorff  $\alpha$  (Krippendorff, 2004)), they can code the rest of the sample independently.

**Codebook Creation** Our codebook was developed by drawing from the literature on framing gun violence, described earlier, as well as from a preliminary analysis of the data. This resulted in 9 frames, including both generic: “Politics”, “Public opinion”, “Society/Culture”, and “Economic consequences” and issue-specific: “2nd Amendment” (Gun Rights), “Gun control/regulation”, “Mental health”, “School/Public space safety”, and “Race/Ethnicity”.

**Unit of Annotation** We choose our unit of annotation to be a news headline for several reasons. Firstly, psychologists have long argued that first impressions are lasting impressions (Digiro-lamo and Hintzman, 1997). This thesis applies to news reading behavior as well. Media framing researchers often identify and measure frames in news headlines (e.g., (Bleich et al., 2015; Trimble and Sampert, 2004), which are seen by the audience first and can determine the perception of the text that follows (Tankard Jr, 2001). As Pan and Kosicki (1993) suggests, a headline is “the most salient cue to activate certain semantically related concepts in readers minds; it is thus the most powerful framing device of the syntactical structure”.

Secondly, the analysis of news headlines became more relevant in the emerging (i.e. digital) media environment where a large portion of people only read headlines but nothing else (Ga-

bielkov et al., 2016). Further, driven by the attention economy, many online media even use news headlines as “clickbait”, presenting sensational but misleading information that deviates from the content included in the actual news story (Chen et al., 2015). That is, a news story may be framed differently in its headline and the rest of the article. In cases like this, research shows that even reading through the article cannot necessarily correct the headlines misdirection (Ecker et al., 2014). Taken together, detecting frames through news headlines provides the most direct clue to the potential influence of the news coverage.

**Annotation Process** Two communication graduate students were recruited to annotate a sample of the collected news articles. They were instructed to first determine whether the news headline was relevant to gun violence in the United States. If yes, they were asked to identify up to two dominant frames in the headline. They were trained on the codebook during the training sessions. In the first training session, the students were given a 100-headline sample to code, and ICR was not met. Hence, a second training session was held to further clarify the codebook and resolve any confusion. The students coded another 100-headline sample, for which ICR was met on all variables: relevance (99% agreement, 0.97  $\alpha$ ), frame A (94.10% agreement, 0.90  $\alpha$ ) and frame B (96.04% agreement, 0.82  $\alpha$ ). Following QCA, once the ICR was met, one student continued to code another 2,790 news headlines, resulting in a total of 2,990 annotated news headlines.

### 3.3 Dataset Properties

GVFC includes 2,990 news headlines, 1,300 of which are annotated as relevant to the gun violence issue in the United States. Out of the relevant headlines, only 319 are found to have 2 frames. Examples of headlines with 2 frames are “It’s Time to Hand the Mic to Gun Owners”, annotated with “Public opinion” (frame A) and “2nd Amendment” (frame B); and “Trevor Noah: ‘The Second Amendment Is Not Intended for Black People’”, annotated with “2nd Amendment” (frame A) and “Race/Ethnicity” (frame B).

We use frame A annotations to train our frame classification model but find that our model also identifies some of frame B annotations in its top predictions (Section 4.2). Table 1 shows frame A class distribution in GVFC that reflects the varying

coverage of different frames in the U.S. media.

## 4 Experiments

We use the most recent method for automatic news frame detection (Field et al., 2018) as one of our baselines. They devise a word-based method for detecting the frames in English and Russian news. They use MFC to derive a lexicon for each frame  $F$  by computing pointwise mutual information  $I(F, w)$  (Church and Hanks, 1990) for each word  $w$  and each frame  $F$  in the corpus. Each frame  $F$ 's lexicon contains the top 250 words with the highest  $I(F, w)$  for frame  $F$ . A news article has a frame  $F$  if it contains at least 3 instances of a word from  $F$ 's lexicon with the primary frame being the most common frame, based on the number of words from each frame's lexicon in the document. We create lexicons for the 9 frames in our GVFC dataset and use them to compute the primary frames of news headlines.

We also implement LSTM-based neural networks for a more comprehensive evaluation. Long short-term memory (LSTM) is a recurrent neural network (RNN) architecture that is widely used today in text classification tasks. There are plenty of variants from this type of architecture: Gated Recurrent Unit (GRU), Bi-directional LSTM, and Bi-directional GRU. We implement these networks with attention mechanism (Bahdanau et al., 2015) and use 100-dimensional pre-trained Glove embeddings (Pennington et al., 2014) as our initial word representations. We train and evaluate these networks for headlines frame classification with 128 units of RNN cells and one layer of attention mechanism at the end, a batch size of 128 for 2000 steps. We use Adam optimizer with a learning rate of 0.01.

As the results in Table 1 show, Bi-directional GRU with attention achieves the highest accuracy among our baselines. The reason behind this could be the fact that we have a small dataset and GRU needs fewer data points to generalize (Kaiser and Sutskever, 2015; Yin et al., 2017). Furthermore, the attention mechanism and bi-directionality allows for more contextual interpretation of the headlines and better detection of their frames.

### 4.1 News Frame Detection with BERT

Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al., 2018) take this idea of attention and bi-directionality further

by building on the Transformer's encoder model that solely relies on multi-layer self-attention to compute contextual representations of its input, dispensing with any kinds of recurrence (Vaswani et al., 2017). The encoder is composed of a stack of identical layers, where each layer contains a self-attention mechanism, which allows the encoder to look at other words in the input sentence as it encodes the contextual representation of each word in the sentence, and a fully connected feed-forward network. The self-attention mechanism computes three vectors from the embedding of each word in the input sentence: the query  $q$ , key  $k$ , and value  $v$  vectors. It then computes the contextual representation of each word  $w$  in the sentence as the weighted sum of the value vectors of all the words in the sentence, where the weights are the scaled then normalized dot products between  $w$ 's query vector and the key vectors of all the words in the sentence. The weights essentially determine how much focus to place on other parts of the input sentence as the encoder encodes a word at a certain position. Given that the query, key, and value vectors are computed by multiplying the input word embeddings matrix  $X$  with weight matrices learned during training,  $W^Q$ ,  $W^K$ ,  $W^V$ , the self-attention output can be formulated in matrix form as:  $\text{Attention}(Q, K, V) = \text{softmax}(QK^T/\sqrt{d_k})V$  where  $Q = XW^Q$ ,  $K = XW^K$ ,  $V = XW^V$ .

BERT's encoder implements the Transformer's multi-layer self-attention mechanisms and fully utilizes its strength in storing the left and right context of each token by using a "masked language model" (MLM) pre-training objective, inspired by the Cloze task (Taylor, 1953). In its pre-training, BERT randomly masks some of the tokens from its input, and predicts the original vocabulary id of the masked word based only on its context. Unlike left-to-right language model pre-training, the MLM objective enables the representation to fuse the left and the right context, which allows BERT to pre-train a deep bidirectional Transformer representations from unlabeled large text corpora.

We fine-tune the pre-trained BERT-based uncased model on our multiclass frame classification by adding a frame classification layer on top of the model and fine-tune all the parameters end-to-end. Given a headline, BERT tokenizes the headline to tokens based on WordPiece tokens (Wu et al., 2016) and appends a special classification token

Frame Class	# Headlines	Baseline	LSTM	Bi-LSTM	Bi-LSTM w/ Attention	Bi-GRU w/ Attention	BERT
2nd Amendment	38	44.74	23.68	21.05	44.74	26.32	65.79
Gun control/regulation	215	50.23	63.72	66.51	72.09	76.28	84.19
Politics	373	40.48	78.28	77.75	84.18	85.79	89.54
Mental health	65	35.38	50.77	40.00	58.46	60.00	78.46
School/Public space safety	137	39.42	48.91	50.36	54.74	58.39	78.10
Race/Ethnicity	114	67.54	75.44	71.93	84.21	81.28	92.11
Public opinion	237	63.29	70.46	72.15	75.53	77.22	86.08
Society/Culture	41	43.90	24.39	19.51	36.59	21.95	58.54
Economic consequences	80	38.75	45.00	51.25	61.25	60.00	80.00
<b>Overall</b>	1300	48.38	64.37	64.48	72.15	72.76	84.23

Table 1: Class distribution of frame A annotations and micro-accuracies for the baseline (Field et al., 2018), LSTM, bi-directional LSTM, bi-directional LSTM and bi-directional GRU with attention, and our method based on fine tuning BERT.

([CLS]) at the beginning of the headline. We use the final hidden vector  $C \in \mathbb{R}^H$  corresponding to [CLS] as the aggregate representation of the headline that is input to the classification layer (since encoding this token with self-attention effectively includes attention to all the tokens in the headline). The only new parameters are our classification layer weights  $W \in \mathbb{R}^{K \times H}$ , where  $K = 9$ , the number of our frame classes. Given the imbalance in our class distribution, we compute a focal loss (FL) (Lin et al., 2017) that improves our classification performance compared to the standard cross entropy loss. We compute  $FL(p) = -\alpha(1-p)^\gamma \log(p)$ , where  $p \in \mathbb{R}^K$  contains the probabilities of classifying the headline into each of the  $K$  frames i.e.,  $p = \text{softmax}(CW^T)$  and  $\alpha \in \mathbb{R}^K$  contains the weighting factors, which we set for each frame to be its normalized inverse class frequency  $\in [0, 1]$ : the smaller the class, the higher the  $\alpha$  and vice versa, which balances the importance of each class’ examples. The modulating factor:  $(1-p)^\gamma$  in FL down-weights the loss contribution of the easy examples – those that are well classified (i.e. have high  $p_k$ ) – and thus focuses the training on hard-to-classify examples. Following Lin et al. (2017), we use  $\gamma = 2$ .

We train for 10 epochs with a batch size of 4,  $2e-5$  learning rate, and maximum sequence length of 128 tokens. Training and testing on the same stratified folds that we use for all our baselines, we achieve a 5-fold cross validation micro-accuracy of 84.23%. Our method based on BERT significantly outperforms not only the most recent news frame classification baseline, but also some state-of-the-art deep classification models, including bi-directional LSTM/GRU with attention on every frame of our GVFC dataset (Table 1).

We also evaluate our method to classify frames of news headlines in another dataset (MFC). As

MFC Issue	# Headlines	Bi-GRU w/ Attention	BERT
Immigration (I)	7231	40.84	52.38
Tobacco (T)	3959	57.20	67.94
Same-sex (SS)	3842	61.57	71.50
I (top-5 frames)	4175	53.65	67.28
T (top-5 frames)	2759	71.44	82.32
SS (top-5 frames)	2937	74.94	83.07

Table 2: 10-fold cross-validation micro-accuracy on the MFC dataset for our best baseline from previous evaluation, and our model based on BERT.

Table 2 shows, our method significantly outperforms our top-performing baseline, both on the 15-frame classification task and on the top-5 (most frequent) frame classification on all issues: immigration, tobacco, and same-sex marriage. This shows that our method can perform well for detecting frames in headlines in different datasets and across a diverse range of issues.

## 4.2 Discussion

Our results show that fine-tuning on BERT performs well even on a small dataset like GVFC, which agrees with the findings of Devlin et al. (2018) that fine-tuning on BERT’s pre-trained model can lead to large improvements even on very small scale tasks. Part of the reasons may be due to BERT’s deep attention mechanism. Attention mechanism has been shown to be data-efficient and helps improve performance significantly even when the dataset is small (Vinyals et al., 2015). Even adding standard attention improves the accuracy of our LSTM-based baselines significantly (Table 1). BERT’s success can also be traced to its design of bidirectional Transformer that offers richer contextual information. Furthermore, BERT was pre-trained on a large corpus to produce this representation. Fine-tuning on BERT allows us to transfer this contextual knowledge to

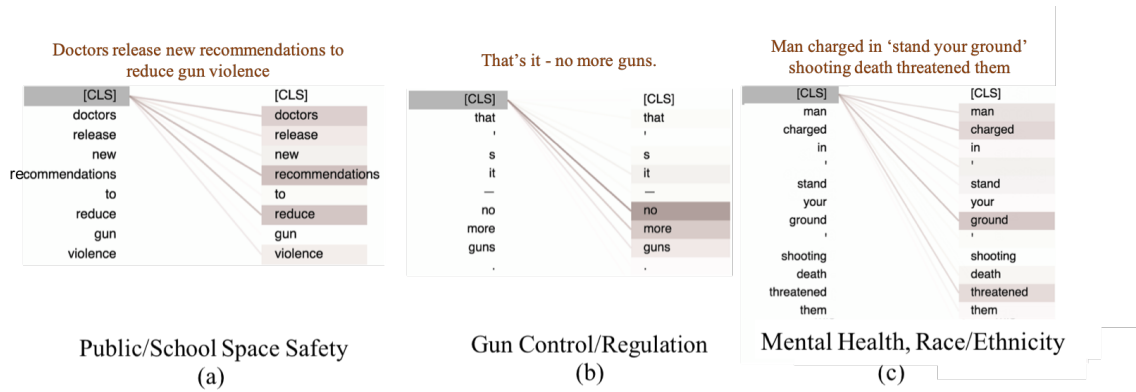


Figure 1: Visualization of our fine-tuned model, the headlines and the predicted frames. The thicker the line, the more attention placed on the token for computing the aggregate i.e., [CLS] representation for the classification

classifying frames in headlines, which are very short compared to the entire news text. The ability to transfer contextual knowledge from a large corpus leads to better representation for these short pieces of texts and better generalization of our model compared to the lexicon-baseline that only relies on word-frame co-occurrences in GVFC.

We use a visualization tool (Vig, 2019) to obtain insights into *what* our fine-tuned model is attending to when making decisions. For example, we observe that pre-training on a large corpus may have helped our model predict the frame “School/Public Space Safety” for the headline: “Doctors release new recommendations to reduce gun violence” by attending to words like “Doctors” and “recommendations” (Figure 1(a)). Although these words do not co-occur frequently with this specific frame in GVFC, they may be related to school/public safety in general. The lexicon model, on the other hand, incorrectly predicts the “Gun control/regulation” frame due to the words “release” and “gun” in the headline.

Because news framing is closely related to journalists’ word choice (Hamborg et al., 2019), we find that on frames such as “Race/Ethnicity,” which has a specific set of keywords that the model can attend to like “black”, “white”, or “anti-semitic”, both our model and the lexicon-baseline perform the best on this frame.

On the other hand, the performance of our model and the baseline differ significantly for generic frames such as “Politics,” whose keywords may overlap with issue-specific frames such as “Gun control/regulation”. Since BERT is pre-trained to take context into consideration, words like “gun”, which appears with all the frames, can have different contextual representation depend-

ing on its context i.e., “gun lobby” vs. “gun permit”. For example, the headline “That’s it – no more guns” is classified correctly by BERT as having “Gun control/regulation” frame by attending to the context “no more” of “guns” (Figure 1(b)).

Also, despite not being trained to predict multiple frames, some of BERT’s predictions of what it believes to be top frames align with that of human experts. There are 319 headlines in GVFC that were annotated with two frames: frame A and B, meaning that the headline is equally likely to belong to either frame. In our experiments, we only train our model with frame A annotations. However, we notice that out of the 319 headlines that have two frames, 164 of them have both frames predicted in the top-2 predictions of our model, showing the potential to fine-tune BERT for multi-label multi-class frame classification, which we will explore in the future. Furthermore, the accuracy of our model on GVFC increases to 87.92% if we consider our model’s prediction to be correct if it predicts either frame for these 319 headlines.

More interestingly, we observe that our model can predict additional frames that may be applicable to the headlines but are not annotated. In Figure 1(c) for example, for the headline “Man charged in ‘stand your ground’ shooting death threatened them”, our model first attends to the word “ground” and then “threatened” and predicts the frame “Race/Ethnicity” and then “Mental health”. Even though this headline was only annotated with the “Mental health” frame (possibly due to the word “threatened” which, in the “Mental Health” description of the annotation codebook, may be referring to an individual’s behaviors that indicate instability, impulsivity, anger, etc.), we believe that in this particular headline the

“Race/Ethnicity” frame is more applicable given the presence of ‘stand your ground’, a legislation that has been shown to have a quantifiable racial bias (Ackermann et al., 2015).

Overall, what we observe from visualizing our model suggests that the model is able to generalize beyond word-frame co-occurrences in the limited annotations by virtue of the contextual knowledge transfer obtained by pre-training on a large corpus.

## 5 Framing Trends, Analysis, and Conclusion

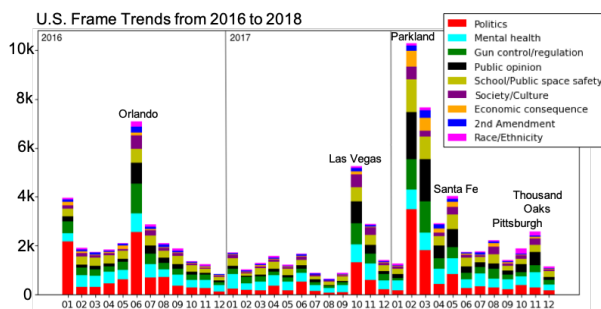


Figure 2: The number of times each frame is represented in new article headlines related to gun violence across the 3-year (per month) period. Some of the peaks represent the deadliest mass shootings in the U.S. since 1949 (CNN Library, 2019).

We used the same search words to retrieve news article headlines from the 21 U.S. news media outlets from 2016 to 2018. To apply our framing analysis, we first train a model to predict whether a news headline is relevant to the issue of U.S. gun violence by fine-tuning BERT-base uncased using the relevance annotations in GVFC. This relevance prediction model achieves a 10-fold cross validation precision of 0.93, 0.95 recall, and 0.94 F-score. We apply this model to find relevant headlines among the 88,470 collected, and apply our frame classification on the relevant headlines.

Several patterns emerged from the framing analysis. It appears that news media of all types have largely politicized the gun violence issue right after each major mass shooting (Figure 2). The focus on party politics, the divide between Democrats and Republicans on the issue dominated the coverage. This finding speaks to the highly polarized political environment in the U.S.

We also observe in Figure 2 that right after the Parkland school shooting in 02/2018, the discussion surrounding “Public opinion”, “School/Public space safety”, and “Economic

consequences” frames increases. The increase in “Public opinion” and “School/Public space safety” frames is due to the growth of student activism in the wake of the shooting. Meanwhile, the increase in the “Economic Consequence” frame is due to the decision of several major companies such as Dick’s Sporting Goods to stop selling assault-style weapons in the wake of the event.

We also observe in Figure 2 that frames that spike during every major shooting event, such as “Politics”/“Public opinion”, are not the most persistent. Their frequency peaked during the month but dropped, often drastically, after. Notably, the “Mental health” frame (the cyan bar) appears to be the most persistent, appearing consistently over time in coverage about gun violence.

Another noticeable cross-media pattern in the U.S. media coverage of the gun violence issue is that the conservative-leaning and neutral media emphasized the mental health of individual gunmen to a greater extent than liberal-leaning media (see the cyan bar representing the “Mental health” frame in the left, center, and right plots of Figure 3). About a quarter of news articles from neutral and conservative-leaning media in 2017 are classified as having the “Mental Health” frame (27% and 22% of the articles respectively). In comparison, only 8% of news articles from liberal-leaning media are classified as having this “Mental Health” frame.

This finding about the conservative media (Figure 3 right) is not surprising because connecting mental illness and mass shooting has been a common stance among pro-gun Republican leaders (i.e., “guns don’t kill people, people kill people”). More surprisingly though (and contrary to the common perception of mainstream media such as NYT, CNN, and CBS being liberal-leaning), our study suggests that these neutral, mainstream media (Figure 3 center) has also largely framed the issue from the aspect of mental health, often *more* than the conservative media, which may indicate conservative media’s strong agenda-setting power in the U.S. media ecosystem.

Media framing scholars have also pointed out the importance of examining what aspects of the story has been left out. In our analysis, the frame of “Society/Culture” – a frame that is important and yet would not attract much web traffic – has not been a focus of gun violence coverage in the U.S. As the results demonstrate (Figure 3), major



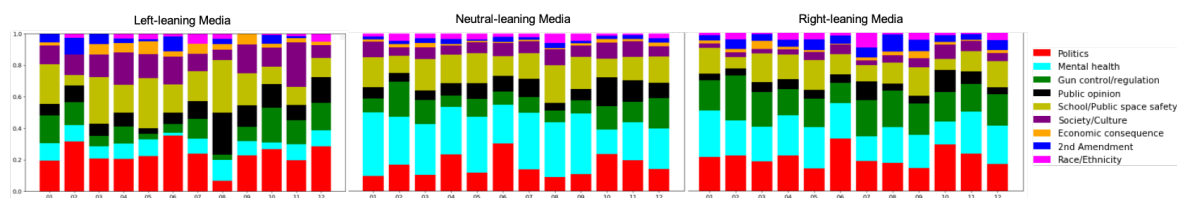


Figure 3: The proportion of each frame occurring in the 2017 news from news sites of different political leanings.

shootings were only able to trigger liberal-leaning media to pay more attention to this frame. 16% of articles from liberal-leaning media in 2017 are classified as having the “Society/Culture” frame, while only 9% of articles from neutral media (and only 5% from conservative-leaning media) are classified as having this “Society/Culture” frame. The lack of framing focus on the underlying cultural/societal issues as well as the aforementioned focus on party politics and strong agenda-setting speak to the status quo of the current U.S. news environment: profit-driven, sensational, and highly partisan.

In conclusion, we have presented in this paper a method for news headline frame classification that achieves state-of-the-art performance. We also release the codebook and a carefully curated Gun Violence Frame Corpus (GVFC) news articles whose headlines have been annotated with their corresponding frames by domain experts. We demonstrate the application of our framing detection to analyze a large corpus of news headlines for framing trends surrounding the U.S. gun violence coverage. We observe interesting findings and believe that frame detection and analysis can potentially be used to gain a deeper understanding of various issues of public affairs. Automatically detected frames in news headlines can also be used to curate more balanced news collections on various issues and perspectives.

**Acknowledgement** We thank the anonymous reviewers for their insightful comments. We thank Mona Jalal and Sha Lai for their help in data collection and figure plotting. This work is supported in part by the U.S. NSF grant 1838193. Any opinions, findings, conclusions, or recommendations expressed here are those of the authors and do not necessarily reflect the view of the sponsor.

## References

Nicole Ackermann, Melody S Goodman, Keon Gilbert, Cassandra Arroyo-Johnson, and Marcello Pagano.

2015. Race, law, and health: Examination of stand your ground and defendant convictions in florida. *Social Science & Medicine*, 142:194–201.

Ad Fontes Media. 2019. The media bias chart, Last accessed on May 31, 2019. <https://www.adfontesmedia.com/intro-to-the-media-bias-chart/>.

Alexa. 2018. The top 500 sites on the web – subcategory “Breaking News.” retrieved from [https://www.alexa.com/topsites/category/top/news/breaking\\_news](https://www.alexa.com/topsites/category/top/news/breaking_news) Last accessed on November 17, 2018.

Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2015. Neural machine translation by jointly learning to align and translate. *ICLR*.

Thomas A Birkland and Regina G Lawrence. 2009. Media framing and policy change after Columbine. *American Behavioral Scientist*, 52(10):1405–1425.

Erik Bleich, Hannah Stonebraker, Hasher Nisar, and Rana Abdelhamid. 2015. Media portrayals of minorities: Muslims in british newspaper headlines, 2001–2012. *Journal of Ethnic and Migration Studies*, 41(6):942–962.

Amber E. Boydston, Dallas Card, Justin Gross, Paul Resnick, and Noah A. Smith. 2014. Tracking the development of media frames within and across policy issues. Technical report, University of California, Davis.

Dallas Card, Amber E Boydston, Justin H Gross, Philip Resnik, and Noah A Smith. 2015. The media frames corpus: Annotations of frames across issues. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, volume 2, pages 438–444.

Yimin Chen, Niall J Conroy, and Victoria L Rubin. 2015. Misleading online content: Recognizing clickbait as false news. In *Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection*, pages 15–19. ACM.

Kenneth Ward Church and Patrick Hanks. 1990. Word association norms, mutual information, and lexicography. *Computational Linguistics*, 16(1):22–29.

- CNN Library. 2019. Deadliest mass shootings in modern us history fast facts. <https://edition.cnn.com/2013/09/16/us/20-deadliest-mass-shootings-in-u-s-history-fast-facts/index.html> Last accessed on May 28th, 2019.
- Ruth DeFoster and Natashia Swalve. 2018. Guns, culture or mental health? Framing mass shootings as a public health crisis. *Health communication*, 33(10):1211–1222.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. BERT: Pre-training of deep bidirectional transformers for language understanding. *CoRR*, abs/1810.04805.
- Gregory J Digirolamo and Douglas L Hintzman. 1997. First impressions are lasting impressions: A primacy effect in memory for repetitions. *Psychonomic Bulletin & Review*, 4(1):121–124.
- Ullrich KH Ecker, Stephan Lewandowsky, Ee Pin Chang, and Rekha Pillai. 2014. The effects of subtle misinformation in news headlines. *Journal of experimental psychology: applied*, 20(4):323.
- Robert M Entman. 1993. Framing: Toward clarification of a fractured paradigm. *Journal of communication*, 43(4):51–58.
- Robert M Entman. 2010. Media framing biases and political power: Explaining slant in news of campaign 2008. *Journalism*, 11(4):389–408.
- Anjalie Field, Doron Kliger, Shuly Wintner, Jennifer Pan, Dan Jurafsky, and Yulia Tsvetkov. 2018. Framing and agenda-setting in Russian news: A computational analysis of intricate political strategies. *arXiv preprint arXiv:1808.09386*.
- Maksym Gabielkov, Arthi Ramachandran, Augustin Chaintreau, and Arnaud Legout. 2016. Social clicks: What and who gets read on twitter? *ACM SIGMETRICS Performance Evaluation Review*, 44(1):179–192.
- Donald P Haider-Markel and Mark R Joslyn. 2001. Gun policy, opinion, tragedy, and blame attribution: The conditional influence of issue frames. *Journal of Politics*, 63(2):520–543.
- Felix Hamborg, Anastasia Zhukova, and Bela Gipp. 2019. Illegal aliens or undocumented immigrants? Towards the automated identification of bias by word choice and labeling. Technical report, University of Konstanz, Germany.
- Crimson Hexagon. 2018. [ForSight social media analytics platform](#), Last accessed on November 1, 2018.
- Kathleen Hall Jamieson, Bruce Hardy, and Daniel Romer. 2007. The effectiveness of the press in serving the needs of American democracy. *A republic divided*, pages 21–51.
- Łukasz Kaiser and Ilya Sutskever. 2015. Neural gpu learn algorithms. *arXiv preprint arXiv:1511.08228*.
- Klaus Krippendorff. 2004. Content analysis: An introduction to its methodology.
- Regina G Lawrence and Thomas A Birkland. 2004. Guns, Hollywood, and school safety: Defining the school-shooting problem across public arenas. *Social Science Quarterly*, 85(5):1193–1207.
- Patricia Leavy and Kathryn P Maloney. 2009. American reporting of school violence and people like us: A comparison of newspaper coverage of the Columbine and Red Lake school shootings. *Critical Sociology*, 35(2):273–292.
- Matthew S Levendusky. 2013. Why do partisan media polarize viewers? *American Journal of Political Science*, 57(3):611–623.
- Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. 2017. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988.
- Emma E McGinty, Daniel W Webster, and Colleen L Barry. 2013. Effects of news media messages about mass shootings on attitudes toward persons with serious mental illness and public support for gun control policies. *American Journal of Psychiatry*, 170(5):494–501.
- Emma E McGinty, Daniel W Webster, Marian Jarlenski, and Colleen L Barry. 2014. News media framing of serious mental illness and gun violence in the United States, 1997–2012. *American Journal of Public Health*, 104(3):406–413.
- MediaCloud. 2018. <https://mediacloud.org>. Last accessed on November 15, 2018.
- Amy Mitchell, Jeffrey Gottfried, Jocelyn Kiley, and Katerina Eva Matsa. 2014. Political polarization & media habits. *Pew Research Center*, 21.
- Nona Naderi and Graeme Hirst. 2017. Classifying frames at the sentence level in news articles. *Polity*, 9:4–233.
- Viet-An Nguyen, Jordan Boyd-Graber, Philip Resnik, and Kristina Miler. 2015. Tea party in the house: A hierarchical ideal point topic model and its application to republican legislators in the 112th congress. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, volume 1, pages 1438–1448.
- Matthew C Nisbet. 2010. Knowledge into action: Framing the debates over climate change and poverty. In *Doing news framing analysis*, pages 59–99. Routledge.
- Zhongdang Pan and Gerald M Kosicki. 1993. Framing analysis: An approach to news discourse. *Political communication*, 10(1):55–75.

- Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.
- Pew Research Center. 2016. Ideological placement of each sources audience, Last accessed on May 31, 2019. [https://www.pewresearch.org/pj14-10-21\\_mediapolarization-08-2/](https://www.pewresearch.org/pj14-10-21_mediapolarization-08-2/).
- Pew Research Center. 2018a. The partisan divide on political values grows even wider, Last accessed on November 17, 2018. <http://www.people-press.org/2017/10/05/the-partisan-divide-on-political-values-grows-even-wider>.
- Pew Research Center. 2018b. State of the news media methodology, Last accessed on November 17, 2018. <http://www.journalism.org/2018/07/25/state-of-the-news-media-methodology>.
- W Russell Neuman, Marion Just, and Ann Crigler. 1993. [Common knowledge: News and the construction of political meaning](#). *Bibliovault OAI Repository, the University of Chicago Press*, 108.
- Frauke Schnell, Karen Callaghan. 2001. Assessing the democratic debate: How the news media frame elite policy discourse. *Political communication*, 18(2):183–213.
- Holli A Semetko and Patti M Valkenburg. 2000. Framing European politics: A content analysis of press and television news. *Journal of communication*, 50(2):93–109.
- Statista. 2017. Most popular news websites as of August 2019, by unique monthly visitors. <https://www.statista.com/statistics/381569/leading-news-and-media-sites-usa-by-share-of-visits/> Last accessed on November 17, 2018.
- Natalie Jomini Stroud. 2011. *Niche news: The politics of news choice*. Oxford University Press on Demand.
- James W Tankard Jr. 2001. The empirical approach to the study of media framing. In *Framing public life*, pages 111–121. Routledge.
- Wilson L Taylor. 1953. cloze procedure: A new tool for measuring readability. *Journalism Bulletin*, 30(4):415–433.
- Linda Trimble and Shannon Sampert. 2004. Who’s in the game? the framing of election 2000 by the globe and mail and the national post. *Canadian Journal of Political Science/Revue canadienne de science politique*, 37(1):51–71.
- Oren Tsur, Dan Calacci, and David Lazer. 2015. A frame of mind: Using statistical models for detection of framing and agenda setting campaigns. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, volume 1, pages 1629–1638.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008.
- Jesse Vig. 2019. Visualizing attention in transformer-based language models. *arXiv preprint arXiv:1904.02679*.
- Oriol Vinyals, Łukasz Kaiser, Terry Koo, Petrov Slav, Sutskever Ilya, and Geoffrey Hinton. 2015. Grammar as a foreign language. In *Proc. of NIPS*.
- Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, et al. 2016. Google’s neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*.
- Wenpeng Yin, Katharina Kann, Mo Yu, and Hinrich Schütze. 2017. Comparative study of cnn and rnn for natural language processing. *arXiv preprint arXiv:1702.01923*.