DocNet: Semantic Structure in Inductive Bias Detection Models

Jessica Zhu
jeszhu@umd.edu
University of Maryland
College Park, MD, USA

Michel Cukier
mcukier@umd.edu
University of Maryland
College Park, MD, USA

Iain Cruickshank
icruicks@andrew.cmu.edu
Carnegie Mellon University
Pittsburgh, PA, USA

Abstract

News will be biased so long as people have opinions. As social media becomes the primary entry point for news and partisan differences increase, it is increasingly important for informed citizens to be able to recognize bias. If people are aware of the biases of the news they consume, they could take action to avoid polarizing echo chambers. In this paper, we explore an often-overlooked aspect of bias detection in media: the semantic structure of news articles. We present DocNet, a novel, inductive, and low-resource document embedding and political bias detection model. We also demonstrate that the semantic structure of news articles from opposing political sides, as represented in document-level graph embeddings, have significant similarities. DocNet bypasses the need for pre-trained language models, reducing resource dependency while achieving comparable performance. It can be used to advance political bias detection in low-resource environments. Our code and data are made available at: https://github.com/jhzsquared/DocNet__CodeData

CCS Concepts

• **Computing methodologies** → *Machine learning approaches*;
• **Information systems** → **Web searching and information discovery**; **Data encoding and canonicalization**.

Keywords

news, natural language processing, political bias

1 Introduction

The news media ecosystem has transformed dramatically in the past decade. Whereas previously only well-established, professional news sources could easily disseminate information, now any outlet with an internet connection and social media presence can share their information to a broad audience [35]. Social media has allowed not just the proliferation of fake news, but also the exacerbation of polarizing echo chambers. Users increasingly and unwittingly are only exposed to and interact with media that already aligns with their biases [9]. As such, to maintain a well-informed society, readers need to be aware of biased reporting so that they can also

know to seek out non-partisan and diverse reporting [38]. Resources for automatic article-level bias detection must be readily available.

Bias detection remains an inexact practice, exacerbated not only by its inherently subjective nature but also by the many different forms it can manifest as [35]. Various datasets have been developed over the past five years that have increased research of biases at the word, sentence, article, and document level [7, 14, 19, 40, 43, 47, 50]. These automatic bias detection models have historically relied on curated linguistic feature-based characterizations of text [19, 41]. More recent studies have sought to use generalizable and computationally expensive embeddings through deep learning-based methods, like contrastive learning [20], pre-trained language models [2, 22, 24, 32, 39, 44], Long Short-Term Memory (LSTMs) neural networks [24, 36], and Large Language Models (LLMs) [25].

These deep learning methods leverage word context. However, they do not explicitly factor in writing style structure. As such, we propose an inductive, graph-based approach to political bias detection using semantic structure. We construct independent word co-occurrence graphs for each news article and discover that their graph embeddings have comparable performance when used to detect political biases as language model embeddings. Our detection model also statistically outperforms LLM-generated bias predictions. In this paper, our contributions are threefold:

(1) We provide a methodology, DocNet, for a novel, low resource, and inductive document embedding and political bias detection model that is not reliant on pre-trained language models.
(2) We demonstrate the value of semantic structure in bias detection within the media ecosystem.
(3) We identify high semantic structural similarity between news on either side of the political spectrum.

2 Background

Bias detection is complex. Researchers have mapped out 17 different categories of media bias that have different methods and intent [35]. Bias detection in the media is further complicated by high text re-use in published news articles [4] and the many different frames available to discuss an issue [6]. A large part of widely circulated online news consists of redistributed or edited copies from news agencies. The decisions behind what to publish, how to edit an article, and what text to re-use, result in instances of informational and lexical biases. Informational biases occur when news articles change the surrounding information in order to shape how an issue is framed, and thereby perceived. Lexical bias occurs when polarizing words or clauses are used to illicit specific emotions toward an issue [14]. Political biases in news articles are generated through these shifts in content and structure. Identification of these rhetorical methods in a high text re-use media ecosystem requires isolating minute shifts in otherwise similar articles. Therefore, news

bias detection models require an effective method of characterizing these shifts.

News Bias Detection. Current news bias detection models use pre-trained word embeddings, like GloVe or Word2Vec, and language models, like BERT, to capture biased language in news articles. For instance, Zou et al., [50] construct an article bias detection model using RoBERTa-base and the POLITICS language model, which is a fine-tuned language model for ideology prediction [27]. Truica et al. combine term frequency-inverse document frequency (TF-IDF), and both non-pre-trained and pre-trained language model word embeddings to represent documents for article bias detection [42]. Meanwhile, Chen et al., use the BASIL dataset [14] and predict article-level biases using the sentence-level biases' frequency, position, and sequential order [8]. They find that for the BASIL dataset, lexical information (e.g. through bag of words), was insufficient in detecting article-level bias, while features that characterize the internal structure of an article were more effective. However, these methods are not easily generalizable. They all hinge on either a corpus of meticulously extracted sentence-level and event features, or a suitable pre-trained language model.

Document Embeddings and classification. Article-level news bias detection can also be reframed as a document embedding and classification problem. Methods remain similar within the more generalized document classification space. Most classification models are highly dependent on embeddings from pre-trained language models, like Longformers [3] or expert generated features. Doc2Vec [23] is one of the few methods that requires neither. However, transformer based approaches, where word embeddings from pre-trained language models are aggregated, have largely surpassed Doc2Vec's performance.

Document classification researchers have also created embeddings via graph representations [26, 45]. Much of graph based document classification research has been transductive, where one graph is generated for the entire corpus of documents, to include the "test" set. Word nodes are connected across documents and document nodes are generated for classification [7, 13, 28, 48, 49]. Transductive models, however, are not representative of reality and are not computationally efficient. They require test documents during training and must be retrained before being applied to new cases.

InducT-GCN is one of the few inductive document classification models discussed in literature [46]. It also creates one graph for the entire training corpus using document node vectors and word node vectors, but can update new nodes during testing. Word-word edges are connected by their pointwise mutual information (PMI) across documents. Word-doc edges are connected using TF-IDF. They then train a graph convolutional network (GCN) model to learn the word embeddings and conduct classification. These word embeddings and learned weights are used to generate predictions on batches of test documents.

However, researchers have found that inductive document-level models struggle with longer documents due to their dense graphs [30]. This is apparent in the case of Induct-GCN, where while still outperforming baselines, it has a distinctly worse performance on longer document datasets, like 20 News Group and Ohsumed, than for shorter ones, like R52 [46].

We address these gaps through our novel, proposed document embedding and classification approach, DocNet. It is inductive, and therefore easily generalizable to new data. It also performs well on news articles bias detection. DocNet only uses the semantic structure in text rather than context dependent comparisons common in the aforementioned studies.

3 Method

In the following section, we detail our experimental methodology. The overall pipeline for our inductive bias detection methodology has four procedures, 1) Dataset collection and preprocessing, 2) Graph construction, 3) Unsupervised embedding generation and 4) Supervised model training. The DocNet pipeline and its experiments are summarized in Figure 1.
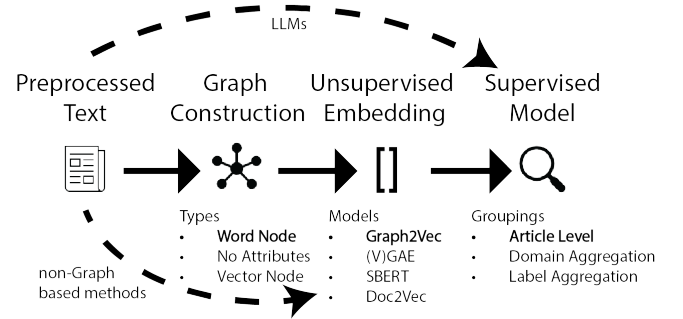
Figure 1: Pipeline of bias detection modeling experiments. The DocNet methodology is in bold. Dotted arrows indicate methods that skip certain steps.

3.1 Datasets

Our analysis uses four datasets, three of which are centered around partisan topics: withdrawal from Afghanistan, the Oath Keepers and the military vaccine mandate. The fourth is the BASIL benchmark dataset [14]. For brevity we will refer to them as the AFG, OATH, VAX, and BASIL datasets, respectively. Each article is cleaned and lemmatized using the code shared.[1]

Table 1 summarizes key characteristics of these four datasets. Tables 2 and 3 list the ten most and least frequent domains in the cleaned topic aligned datasets.

Topic Aligned Datasets. We introduce three topic-based datasets, the AFG, Oath, and VAX news articles, and have made them publicly available. We chose these three topics for experimentation due to the significant attention, controversy, and corresponding media coverage surrounding each. Each topic's respective articles were collected by first pulling Tweets from February 1st, 2022 until December 29nd, 2022. These Tweets were pulled using a keyword search of relevant terms via Twitter's then available Researcher API. Table 4 displays the collection terms used.

Per the methodology of other works [10, 11], we extracted URLs from Tweets and filtered out social media site URLs to ensure only

---

[1]https://github.com/jhzsquared/DocNet__CodeData

Table 1: Summary of datasets (Note: "Far-Left" labels do not exist in this data. "Sentences" is shortened to "sent")

Dataset	AFG	OATH	VAX	BASIL
# of Articles	5490	1965	1367	300
# of Domains	746	417	400	3
Sent/Article	39.3	40.4	37.0	26.6
Words/Sent	23.0	24.1	23.4	28.1
%Left	8.3	20.5	3.9	28.0
%Left-Center	40.8	48.0	23.6	00.7
%Center	20.2	23.8	26.1	46.0
%Right-Center	14.9	4.4	12.3	0.0
%Right	15.4	3.2	26.1	25.3
% Far Right	0.3	0.2	3.7	0.0

Table 2: Top 10 most frequent domains

Domain	Label
apnews.com	left-center
theepochtimes.com	right
washingtonpost.com	left-center
dailymail.co.uk	right
cnn.com	left
reuters.com	center
foxnews.com	right
msn.com	left-center
news.yahoo.com	left-center
theguardian.com	left-center

Table 3: 10 Least frequent domains

Domain	Label
mb.com.ph	left-center
idahopress.com	left-center
icrw.org	left-center
statesville.com	center
i24news.tv	center
humansarefree.com	right
humanevents.com	right
steadfastandloyal.com	right
hrc.org	left
hoover.org	right-center

external sites remained. We next scraped the textual content of the websites that allowed automated scraping. This methodology results in a variety of news sources and articles that are relevant to discussions on these partisan topics [12]. We align the collected domains to labels from a dataset of news domains and their biases compiled by Cruickshank and Carley [11]. Domains are the publishing sites that news articles are scraped from. The term domain is used in lieu of news sources, as articles are often shared from distributing news agencies, like the Associated Press, to multiple sites (i.e. domains).

Table 4: Collection terms for each dataset

Dataset	Collection Terms
AFG	{afghanistan, taliban, kabul, bagram U.S. withdrawal, afghan}
OATH	{oath keepers}
VAX	{vaccine, vax}

A weak bias label (far left, left, left-center, center, right-center, right, far right) for each article is assigned based upon its domain's bias label. Our usage of domain labels as weak labels for article bias follows the assumption that for highly contentious topics, the bias of a news article and the events that are discussed are highly correlated with the perspective of its publisher, or domain [50]. We further validate the reliability of the domain bias labels by comparing them with the domain bias labels published on Adfontes, as of 10 March 2023[2]. From a Pearson correlation test of news domain bias labels using an ordinal encoding, we validate that our labels are highly correlated with Adfontes' with a correlation coefficient of .79 and p-value $< e^{-80}$. This is in spite of the three year difference in times at which these domain bias labels were collected (2020 vs 2023) and the likely drift in media bias over time. This test increases confidence that the domain bias labels we use have minimal biases themselves, as both data sources report similar labels.

BASIL. We further validate our approach with a dataset that has human annotated article and domain-level biases, the BASIL dataset [14]. The BASIL dataset is composed of 300 articles from 2010 to 2019 that cover 100 events, with one article each from Fox News, New York Times, and Huffington Post. These domains are considered to be right, center, and left, respectively.

3.2 Graph Construction

In order to represent each article's semantic structure, the syntactical patterns used to create meaning, we transform each news article into its own graph. Each graph is an undirected word (or lemma) co-occurrence network. An edge is drawn between co-occurring words in a sentence, where the weight of the edge is the positive pointwise mutual information (PPMI), Equation 1. The edge is removed if there is a non-positive PMI in order to narrow down edges to connections with a strong association.

$$PMI = \frac{log_2(P_{w_1 w_2}/(P_{w_1} * P_{w_2}))}{-log_2(P_{w_1 w_2})})$$

$$P_x = \frac{\text{Frequency of } x \text{ in document}}{\text{Total Unigrams or Bigrams}} \quad (1)$$

w_1, w_2 are the words that construct edge (w_1, w_2)

We construct three variations of these article graphs:

(1) Base: No node attributes are provided.
(2) Word Nodes: The node attributes are the unique word. This is the graph used in DocNet.

---

[2]https://adfontesmedia.com/

(3) Vector Nodes: The node attribute is the word embedding vector per Spacy's en_core_web_lg model[3] [16] and concatenated with the average sentiment of the given lemma across all occurrences in the article per NLTK VADER's Sentiment Intensity Analyzer [17].

3.3 Unsupervised Embeddings

We then embed each article using unsupervised models. We experiment with 5 different embedding methods, two of which embed the graph representation of each article, two of which embed the article's cleaned text, and one hybrid approach. Except for SBERT, which is pre-trained, these embedding models are inductively trained with the same 80%/20% train/test splits. We do not conduct hyperparameter tuning. Instead, we seek to align parameters, like epochs and dimension size, across models for more equitable comparison, and otherwise use default package hyperparameters. No notable performance improvements were observed during initial experiments to warrant the additional computational cost of hyperparameter tuning.

Graph2Vec. We train three Graph2Vec models for each graph variation using the KarateClub package [37]. Graph2Vec conducts Weisfeiler-Lehman (WL) hashing to generate graph-level features that are then passed as the "document" in Doc2Vec [23]. In the base graph's case, the hashing step is conducted using node degrees, whereas the designated node attribute is used otherwise. The PPMI scores are not factored into the model as Graph2Vec does not currently support edge attributes. We set a dimension size of 128, 50 epochs, and 2 Weisfeiler-Lehman iterations.

Graph Auto-Encoders. We also train graph auto-encoder models ("GAE") and variational graph auto-encoder models ("VGAE") [21] on the Vector Node graph representation of the text, both with and without the sentiment vector ("sent"). Both graph auto-encoders are constructed using three modules of graph convolution, graph normalization, and rectified linear unit (ReLU) layers in the encoder, and an inner product decoder. In order to have a comparable embedding size and epochs to the other models, we set output feature size to be 128. We use an Adam optimizer with a starting learning rate of .1 and a scheduler that reduces the learning rate when the loss plateaus after a minimum of five epochs and maximum of fifty epochs.

Doc2Vec. As a comparison to the graph-based approaches, we use Doc2Vec [23], using both the distributed memory ("Dm") and the distributed bag of words ("Dbow") representations. We set a window size of 10, initial learning rate of .05, 50 epochs, minimum word count to be 3, negative sample of 5, and a vector size of 128.

Doc2Vec and Graph2Vec. In order to measure if graph representations capture anything different from document representations, we concatenate the separately trained Graph2Vec and Doc2Vec embeddings into a 256 dimension sized vector. We combine the two variants of Doc2Vec with the three variants of Graph2Vec for a total of six different embedding models.

---

[3]Version 3.4.1

Sentence BERT. As a final unsupervised embedding comparison, we use Sentence BERT (SBERT) [33]. For each sentence in an article, we pass it through the pre-trained "All-Distilroberta-V1" encoder to map it to a default 768 dimensional vector space. This language model is a top performing sentencing language model of modest size. Each article is represented by the average of all of its sentence embeddings.

3.4 Bias Detection

We explore the value of these embeddings in political bias detection using a naive multinomial logistic regression model on the same train/test splits as the unsupervised embeddings. Along with training a supervised model at the article-level, we also experiment with domain aggregation, and bias label aggregation. These variations allow us to explore the level of granularity of the bias information captured in the various embeddings. It also allows us to interrogate the potential for similarities between opposing viewpoints in partisan discussions. The results from our pipeline are also compared with those from LLMs and a naive baseline.

Domain Aggregation. We explore domain-level bias detection models due to the weak article bias labels used in the AFG, OATH, and VAX datasets. In the topic aligned datasets, there is a mixed number of articles between domains. As a result, we de-duplicate domains for training a news domain bias detector using various methods of aggregating article embeddings. In other words, the models are trained on a combined representation of each domain's article embeddings.

We experiment with three methods of aggregating articles for each domain. First, we aggregate articles in a given domain by taking the arithmetic mean of the embeddings ("Mean"). We also explore the effect of normalizing the embeddings by its original topic. We subtract the average embeddings within a given topic from the embedding of each article and take the average ("Topic Diff-Avg", Equation 2), and for another method, take the Frobenius norm ("Topic Diff-Norm", Equation 3). The latter two aggregation techniques aim to remove any topic based skew in the data.

$$T_t = \frac{1}{n_t} \sum_{i=1}^{n_t} x_{ijt}$$

$$\hat{x}_{ijt} = x_{ijt} - T_t \qquad (2)$$

$$\text{avgTopicDiff: } d_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \sum_{t=1}^{n_t} \hat{x}_{ijt}$$

$$\text{froNormTopicDiff: } d_j = \sqrt{\sum_{i=1}^{n_j} |\hat{x}_{ijt}|^2} \qquad (3)$$

d_j = embedding of domain j,
x_{ijt} = embedding of article i from domain j and topic t,
n = number of articles, and
T_t = average topic embedding of topic t

For the BASIL dataset, our domain-level models predict the domain bias labels for each article without aggregating embeddings since there are only three different news domains. This increases the risk of the model capturing spurious characteristics, but in

reviewing our cleaning procedure, we believe the risk of domain specific text cues being present is sufficiently mitigated.

Label Aggregation. We train three separate inductive models using varied bias label groupings 1) "Full": the original bias categories, 2) "LCR": Re-classifying mildly or extremely partisan articles into right or left, for three nominally encoded bias categories (right, center, left) and 3) "Binary": Re-classifying into biased or unbiased, where left/right-center are both considered to be unbiased.

After initial predictions are generated, we regroup the "Full" category predictions into"LCR" and "Binary" to more accurately compare the accuracy of models with different numbers of categories.

Large Language Models. As a final comparison, we conduct document level bias detection using pre-trained large language models. We experiment with both OpenAI's GPT Turbo 3.5, which is optimized for chat completion but has also demonstrated strong performance for general purpose tasks [29], and Meta's Llama 3 8B Instruct Model, which is tuned for instruction tasks [1]. We pass the following zero-shot prompt template to generate bias predictions: "Please classify the following news article by its political bias. Please only classify the article as {bias labels}, and return no other text. title: {title} article: {article} bias: ". We do not conduct domain-level predictions using the large language models.

Naive Baseline. We use a Zero Rule baseline for a naive model standard, where the test set predictions are all labeled as the majority class in the training set.

4 Results and Discussion

We ran 1284 experiments to train multiple unsupervised embeddings and supervised learning models over the various configurations and each of the four datasets. We also train a topic agnostic dataset that merges the AFG, OATH, and VAX datasets. The experiments are trained on 1 node of an NVIDIA DGX A100 system. Each graph autoencoder embedding model takes at most 2 GPU hours to train one model, while each Graph2Vec embedding model takes at most 3.5 CPU minutes. In order to control for different configurations, we formally evaluate the difference between configurations and embeddings using a one-sided Wilcoxan signed rank test on the accuracy and macro F-1 scores. Since each pairwise configuration uses the same test sets but are trained independently, the necessary assumptions for a valid hypothesis test are satisfied. This test allows us to determine if there is a paired difference between two methods when all other configurations are held constant.

Through this hypothesis test, we confirm that the DocNet embeddings outperform the naive baseline, with p-values close to 0. They have an average .21 increase, or a 120% improvement in macro F-1 scores over the baseline. In addition, DocNet outperforms all the embedding models, except SBERT. However, in 30% of the configurations, it also outperformed the SBERT embeddings. For the sake of brevity, we display the average and standard deviation of scores across configurations in Table 5.

Label Aggregation. As expected, there is a large drop in performance as the number of classes increase, with the models trained

Table 5: Accuracy and macro F-1 by embedding configurations. Avg (std)

Model	Accuracy	Macro F-1
Naive Baseline	0.522 (.159)	0.246 (.133)
LLMs		
GPT 3.5 Turbo	0.417 (.164)	0.349 (.177)
Llama 3-8B	*0.454 (.225)*	*0.378 (.204)*
Text Embeddings		
SBERT	**0.607 (.154)**	**0.542 (.162)**
Doc2Vec: Dm	0.539 (.170)	0.472 (.170)
Doc2Vec: Dbow	0.574 (.170)	0.486 (.170)
Graph Embeddings		
VGAE	0.470 (.155)	0.362 (.135)
GAE	0.538 (.146)	0.425 (.138)
VGAE: w/ sent	0.475 (.143)	0.358 (.128)
GAE: w/sent	0.524 (.146)	0.391 (.135)
Graph2Vec: Base	0.498(.169)	0.341 (.137)
Graph2Vec: Vector Node	0.589 (.154)	0.501 (.161)
Graph2Vec: Word Node (DocNet)	*0.592 (.151)*	*0.507 (.157)*
Hybrid: Doc2Vec+Graph2Vec		
Dbow + Base	0.559 (.169)	0.457 (.170)
Dm + Base	0.544 (.178)	0.467 (.184)
Dbow + DocNet	0.583 (.164)	0.501 (.167)
Dbow + Vector Node	0.585 (.163)	0.501 (.165)
Dm + Vector Node	0.572 (.157)	0.509 (.160)
Dm + DocNet	*0.570 (.164)*	*0.511 (.167)*

for fine grained bias classification contributing to the lower overall average scores. The results from models grouped by various label bins are in Table 6. We also find that after regrouping the predictions from models trained using "Full" labels, there is not a statistically significant difference in macro F-1 in comparison to models trained using the "LCR" or "Binary" labels, with p-values of .91 and .36, respectively. This demonstrates that while the "Full" multiclass models have low scores, they are able to extract characteristics of bias in news just as effectively as the models trained on fewer categories. A one-vs rest model training approach would be needed to validate if more granular bias detection is possible with these embeddings.

Domain Aggregation. The results for the different domain aggregation comparisons are provided in Table 7. For the BASIL dataset, the embeddings were more accurately classified into the domain-level labels, though the article-level predictions were still significantly better than the naive baseline, with a p-value near 0 for all metrics. The method of aggregating article embeddings also makes a significant difference on classification performance, such that domain models using Topic Diff-Norm performs the worst. The mean of article embeddings across samples of a domain has the best performance in domain-level bias detection. More research is needed to identify if topic information is not significant in domain-level bias detection or if a better way of extracting topic-level information is warranted.

Table 6: Accuracy and macro F-1 by label aggregation method versus the naive baseline. "Models" refers to the pooled results across all configurations. Avg (std)

Label Aggregation		Accuracy	Macro F-1
Full			
	Models	0.404 (.102)	0.294 (.111)
	Baseline	0.380 (.075)	0.105 (.038)
LCR			
	Models	0.518 (.102)	0.465 (.102)
	Baseline	0.470 (.067)	0.2167 (.036)
Binary			
	Models	**0.722 (.099)**	**0.608 (.118)**
	Baseline	0.717 (.069)	0.417 (.024)

Table 7: Accuracy and macro F-1 by domain aggregation method and dataset subsets. Avg (std)

Domain Aggregation	Accuracy	Macro F-1
No Aggregation	0.565 (.165)	0.453 (.167)
Mean	**.563 (.168)**	**0.490 (.175)**
Topic Diff-Norm	0.500 (.159)	0.397 (.153)
Topic Diff-Avg	0.556 (.163)	0.483 (.167)
BASIL: Article-Level	0.372 (.111)	0.332(.112)
BASIL: Domain-Level	0.588 (.190)	0.544(.196)
Dataset Specific Models		
BASIL	0.480 (.189)	0.438 (.191)
Full	0.520 (.173)	0.445 (.174)
AFG	0.549 (.157)	0.447 (.143)
VAX	0.510 (.148)	0.451 (.166)
OATH	**0.560 (.161)**	**0.461 (.160)**

4.1 Embedding Model

The average mean and standard deviation of scores given different embedding models are in Table 5. Through the paired rank test on the full set of experiments, the GPT, Llama, graph autoencoders, and Graph2Vec with no features (i.e. the base graph) perform the worst, with no statistically significant performance difference between them. The unsupervised Graph2Vec based embeddings with node attributes are able to capture characteristics of biased text. These embeddings have performance similar to, and with the Word Node graphs, 30% of the time, better than SBERT. However, on average SBERT still outperforms all other embeddings. Graph2Vec models with node attributes (both with Word Nodes and Vector Nodes) outperform all Doc2Vec embedding models with p-values less than .01. When Graph2Vec embeddings with node features are combined with Doc2Vec, a statistically significant improvement (p-values less than .01) from the original Doc2Vec based model occurs. For instance, Graph2Vec with Word Nodes (i.e. DocNet) concatenated with Doc2Vec with distributed memory consistently outperforms embeddings using just Doc2Vec, as observed in their average scores

in Table 5. We do not find a statistical difference between embedding performance based on the type of node attribute used in Graph2Vec, Word Nodes or Vector Nodes.

Although our graph-based embeddings do not consistently outperform SBERT, these results demonstrate the value of including semantic structure in bias classification. Unlike with SBERT, language models and word embeddings are unnecessary in our DocNet methodology. As demonstrated by the performance of the Word Node based embeddings, as long as there is a unique node identifier that indirectly relates unique graphs, the article graph embeddings are comparable to SBERT. This indirect connection between news articles is necessary and sufficient to have performance that is comparable to pre-trained language models. Without factoring in edge attributes, pre-trained word embeddings, or word frequency, our document networks are able to successfully embed information for bias detection.

In contrast, even though our computationally expensive graph autoencoder models pass both edge and node attributes in generating document embeddings, they under perform in comparison to using Graph2Vec. A simple approximate graph isomorphism comparison through WL-hashing captures enough news article information for good bias predictions. The effectiveness of the Word Node graphs validates the value of characterizing semantic structure using our DocNet pipeline.

4.2 DocNet Characteristics Case Study

In the following section, we investigate if DocNet's Word Node graph embeddings align with standard graph features.

We obtain descriptive graph metrics for every article and calculate their Pearson correlation with an ordinal encoding of the bias labels as shown in Figures 2 and 3. There is no strong correlation between any graph metrics and bias labels. We then train a logistic regression model using these metrics to predict binary bias labels. These models have an average .21 lower Macro F-1 on the topic aligned datasets and a .12 lower Macro F-1 on the domain-level BASIL predictions compared to the corresponding Graph2Vec with Word Node models. This analysis demonstrates that while the semantic structure of articles holds bias information, our proposed graph embeddings provide additional information over curated descriptive graph metrics.

We also plot graphs from top performing binary Node Word based embedding models to manually inspect for any consistent characteristics or discrepancies in the data labels. We highlight graphs from the domain label BASIL model, and no domain aggregation VAX, and OATH datasets in Figures 4, 5 and 6, respectively. These are the article graphs with the highest probability incorrect and correct predictions for both bias labels. They are graphed using a spring force-directed algorithm [15].

We observe that in the graphs of BASIL articles, Figure 4, the graphs for the articles that were predicted to be unbiased (a and c) have visually sparse layouts. Meanwhile, articles predicted to be biased (b and d) appear to be denser. Upon returning to the original data, we find that the news article from Figure 4c is actually labeled at the article-level to be unbiased. The original article label for Figure 4d also is unbiased. However Figure 4d's article contains phrases like "Donald J. Trump leads a familiar refrain" and "Mr.
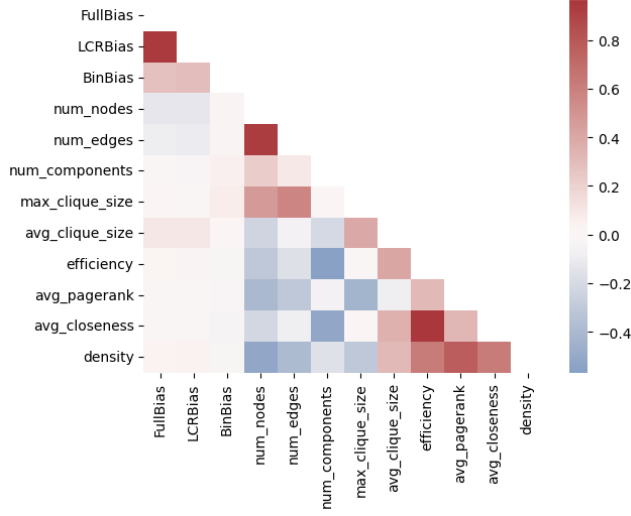
Figure 2: Correlation plot of graph metrics and bias from topic aligned dataset
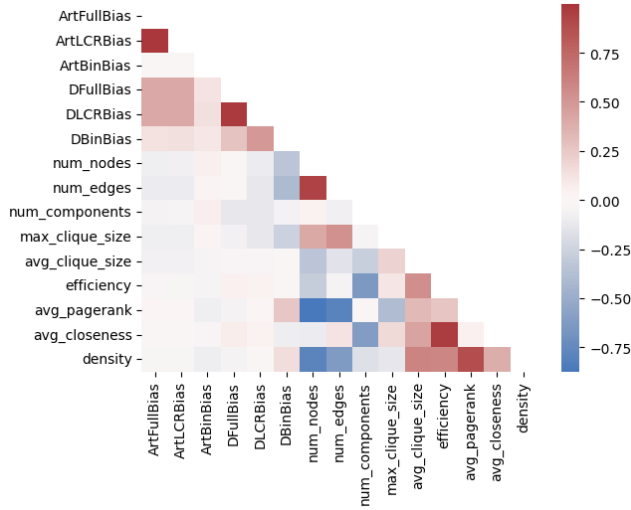
Figure 3: Correlation plot of graph metrics and bias from BASIL (Art is article bias, D is domain bias)

Trump, who has been criticized for lacing his campaign speeches with profanity". These phrases may be considered to be inflammatory given its usage of phrases like "familiar refrain" and words like "lacing" and "profanity" [34]. This example demonstrates the continued difficulty in developing reliable article bias labels for an inherently subjective application. As the BASIL dataset only has three domains, these visual patterns may be from a spurious correlation.

To interrogate these patterns further, we look at the Node Word graphs of the articles using the article-level models of the VAX and OATH datasets, as shown in Figures 5 and 6. As is apparent in these figures, no consistent visual pattern exists between the articles with
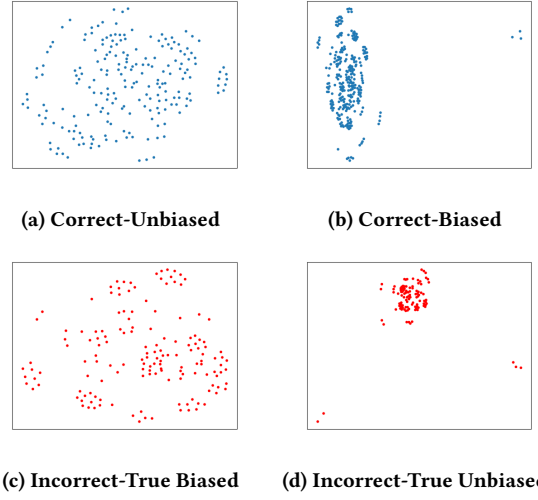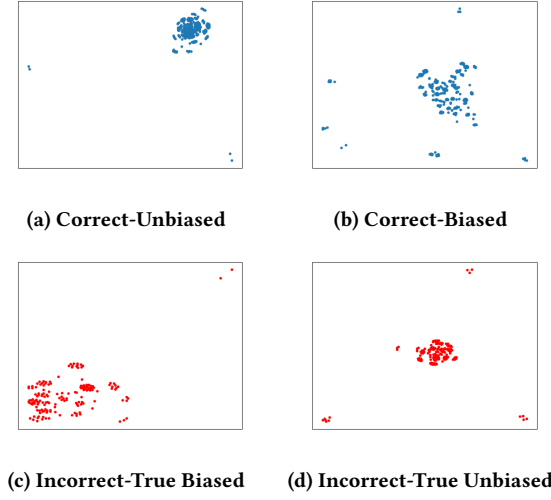

Figure 4: Graphs of articles with high probability predictions from BASIL. Predictions from DocNet (Graph2Vec: Word Nodes) at the domain-level using binary labels (Macro F-1 = .82). Source domains: a) NYTimes, b) Fox, c) Huffington Post d) NYTimes

biased versus unbiased predictions. This reinforces our findings from the graph metric analysis that a graph embedding approach is needed to capture the complex semantic structure characteristics of news articles.
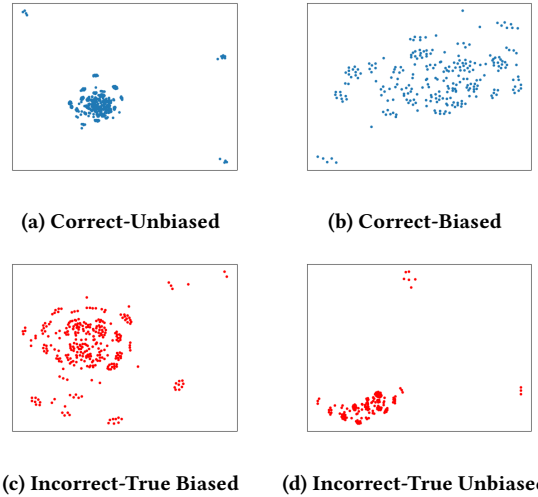
Figure 5: Graphs of articles with high probability predictions from the VAX dataset. Predictions from DocNet (Graph2Vec: Word Nodes) using binary labels (Macro F-1 = .75). Source domains are: a) ABC, b) AFN, c) Toronto Sun d) MSN

Since the topic aligned datasets have significantly greater variety in domains in comparison to the BASIL dataset, the possibility

Figure 6: Graphs of articles with high probability predictions from the OATH dataset. Predictions from Doc2Vec: Dbow + Graph2Vec: Word Nodes using binary labels (Macro F-1 = .76). Source domains are: a) MSN, b) Raw Story, c) Salon d) MSN

the bias detection model learned spurious domain-specific characteristics is reduced. In manual validation of the weak labels from Figures 5 and 6's graphs, we find clear usage of informational bias. The article from Figure 6b ends with the sentence, "Those words are directional instructions on Claymore anti-personal mines"[5], which invokes a sense of incredulity of the actions of the Oath Keepers. Similarly, the AFN article from Figure 5b opens with the title, "Military docs drop vaccine bombshell but 'fact-check' has an explanation" [18]. The usage of "bombshell" and quotation marks around "fact-check" are examples of lexical bias. Both articles align with their domain bias labels. On the other hand, the incorrectly classified OATH MSN article (Figure 6d) is actually originally sourced from CNN. It contains artifacts like "screaming match", "The two continued to cut each other off...", and "chastised". Although its partisan lean is not immediately clear without greater context, these informational and lexical rhetorical decisions lead to the assessment that this article is biased. If the article had been correctly attributed to CNN in our dataset, it would have been labeled as biased, instead of the unbiased, or more specifically left-center, label given to MSN [31]. This inspection demonstrates that while our labeling assumptions falters on borderline biased articles, the generalizability of domain labels to weak article labels for contentious topics largely holds. Our bias detection models may even be more astute than these weak labels.

5 Conclusion and Future Work

Through DocNet, we demonstrate that encoding semantic structure in news articles through word co-occurrence networks can be used for domain and article-level bias detection. We use a logistic regression to learn a separation of the unsupervised embedding space created from these networks. Using DocNet, we also show that biased texts have similar context agnostic semantic structures. These

similarities cross article topics and partisan stances. Through the inclusion of semantic structure in the form of word-co-occurrence networks, our proposed methodology has performance comparable to, and occasionally better than embeddings from transformer based language models.

DocNet successfully uses Graph2Vec with Node Words and a multinomial logistic regression to detect bias. DocNet does not require pre-trained language models to generate word embeddings, nor does it require computationally expensive deep learning techniques. The most computationally time consuming portion of Doc-Net is the initial graph construction. However, this can be parallelized across a corpus of documents for time savings. It is an inductive approach, allowing for inexpensive and generalizable predictions on new datasets.

Through the varied groupings of bias labels, we find that our approach struggles with fine-grained labels. This is an observation also made by other recent works [7]. Further research using one-vs-rest classifiers is needed to confirm if semantic structure embeddings can capture fine-grained differences. On the other hand, even with a naive supervised model, semantic structure embeddings can sufficiently characterize coarser bias labels. This also means that right-leaning news articles and left-leaning news articles have significantly more in common, in terms of article structure and writing styles, than either side of the political spectrum may believe. More research is needed to pinpoint exactly what these structural characteristics or writing style biases are. Regardless, we can conclude that the characteristics of partisan rhetoric transcend a particular topic or context.

In future work, it would be valuable to research alternative methods to our graph autoencoders as well as hybrid approaches that can fuse word embeddings with semantic structure. Alternative deep learning graph models, like graph attention or diffusion based approaches [26, 28] could be explored for improved coupling of context and structural information. New graph hashing algorithms that efficiently encode edge attributes could also be developed and used in Graph2Vec. Future developments in inductive graph embedding approaches that efficiently and effectively represent dense networks would likely improve bias detection efforts. In the interim, our DocNet methodology provides a pathway for low-resourced bias detection in news, when computational resources, rigorously labeled datasets, or language models are not available. As an alternative to large language models, these findings may improve accessibility to underrepresented communities. It helps ensure informed news consumption as people recognize the biases in their online media ecosystems.

6 Limitations

As previously mentioned, this research relies on politically contentious topics and does not look into bias detection of issues with less polarization. It also focuses on political biases. More experimentation would be needed to validate its generalizability to other forms of biases. In addition, since the data was collected based on articles shared over Tweets, there may be a skew in our dataset to highlight more inflammatory media with characteristics that draw the attention of Twitter users. However, given the size of the dataset, the variation in domains present, and the consistency in

trends with the manually labeled and mixed topic BASIL dataset, we believe our data to be sufficiently generalizable. Our data should only be used for research purposes. An incorrect bias detection model may lead to misplaced mistrust or over confidence in media. Additionally, this project only analyzes articles in the English language. While our DocNet methodology is not tied to a language, future work is needed to validate its applicability to news articles from other languages.

References

[1] AI@Meta. 2024. Llama 3 Model Card. https://github.com/meta-llama/llama3/blob/main/MODEL_CARD.md unpublished, from https://github.com/meta-llama/llama3/blob/main/MODEL_CARD.md.

[2] Ramy Baly, Georgi Karadzhov, Jisun An, Haewoon Kwak, Yoan Dinkov, Ahmed Ali, James Glass, and Preslav Nakov. 2020. What Was Written vs. Who Read It: News Media Profiling Using Text Analysis and Social Media Context. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Online, 3364–3374. https://doi.org/10.18653/v1/2020.acl-main.308

[3] Iz Beltagy, Matthew E. Peters, and Arman Cohan. 2020. Longformer: The Long-Document Transformer. *CoRR* abs/2004.05150 (2020). arXiv:2004.05150 https://arxiv.org/abs/2004.05150

[4] Jelle Boumans, Damian Trilling, Rens Vliegenthart, and Hajo Boomgaarden. 2018. The Agency Makes the (Online) News World go Round: The Impact of News Agency Content on Print and Online News. *International Journal of Communication* 12, 0 (2018), 1768–1789. https://ijoc.org/index.php/ijoc/article/view/7109

[5] Bob Brigham. 2022. Stewart Rhodes wanted to dress up like the elderly to trick counter protesters into fights: Oath Keeper. *Raw Story* (2022). https://www.rawstory.com/oath-keepers-trial-2658407982/

[6] Dallas Card, Amber E. Boydstun, 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)*, Chengqing Zong and Michael Strube (Eds.). Association for Computational Linguistics, Beijing, China, 438–444. https://doi.org/10.3115/v1/P15-2072

[7] Peter Carragher, Evan M. Williams, and Kathleen M. Carley. 2024. Detection and Discovery of Misinformation Sources Using Attributed Webgraphs. *Proceedings of the International AAAI Conference on Web and Social Media* 18 (May 2024), 214–226. https://doi.org/10.1609/icwsm.v18i1.31309

[8] Wei-Fan Chen, Khalid Al Khatib, Benno Stein, and Henning Wachsmuth. 2020. Detecting Media Bias in News Articles using Gaussian Bias Distributions. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, Trevor Cohn, Yulan He, and Yang Liu (Eds.). Association for Computational Linguistics, Online, 4290–4300. https://doi.org/10.18653/v1/2020.findings-emnlp.383

[9] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. 2021. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences* 118, 9 (2021), e2023301118. https://doi.org/10.1073/pnas.2023301118 arXiv:https://www.pnas.org/doi/pdf/10.1073/pnas.2023301118

[10] Iain J. Cruickshank and Kathleen M. Carley. 2021. Clustering Analysis of Website Usage on Twitter During the COVID-19 Pandemic. In *Information Management and Big Data*, Juan Antonio Lossio-Ventura, Jorge Carlos Valverde-Rebaza, Eduardo Díaz, and Hugo Alatrista-Salas (Eds.). Springer International Publishing, Cham, 384–399.

[11] Iain J. Cruickshank and Kathleen M. Carley. 2021. Clustering Analysis of Website Usage on Twitter During the COVID-19 Pandemic. In *Information Management and Big Data*, Juan Antonio Lossio-Ventura, Jorge Carlos Valverde-Rebaza, Eduardo Díaz, and Hugo Alatrista-Salas (Eds.). Springer International Publishing, Cham, 384–399.

[12] Iain J. Cruickshank, Jessica Zhu, and Nathaniel D. Bastian. 2023. Analysis of Media Writing Style Bias through Text-Embedding Networks. arXiv:2305.13098 [cs.SI]

[13] Yong Dai, Linjun Shou, Ming Gong, Xiaolin Xia, Zhao Kang, Zenglin Xu, and Daxin Jiang. 2022. Graph Fusion Network for Text Classification. *Knowledge-Based Systems* 236 (Jan. 2022), 107659. https://doi.org/10.1016/j.knosys.2021.107659

[14] Lisa Fan, Marshall White, Eva Sharma, Ruisi Su, Prafulla Kumar Choubey, Ruihong Huang, and Lu Wang. 2019. In Plain Sight: Media Bias Through the Lens of Factual Reporting. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Kentaro Inui, Jing Jiang, Vincent Ng, and Xiaojun Wan (Eds.). Association for Computational Linguistics, Hong Kong, China, 6343–6349. https://doi.org/10.18653/v1/D19-1664

[15] Aric Hagberg, Pieter Swart, and Daniel S Chult. 2008. *Exploring network structure, dynamics, and function using NetworkX*. Technical Report. Los Alamos National Lab.(LANL), Los Alamos, NM (United States).

[16] Matthew Honnibal and Ines Montani. 2017. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. (2017). To appear.

[17] C. Hutto and Eric Gilbert. 2014. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media* 8, 1 (May 2014), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

[18] Steve Jordahl and Davis Billy. 2022. Military docs drop vaccine bombshell but 'fact-check' has an explanation. *NY Times* (2022). https://afn.net/medical-health/2022/02/03/military-docs-drop-vaccine-bombshell-but-fact-check-has-an-explanation/

[19] Johannes Kiesel, Maria Mestre, Rishabh Shukla, Emmanuel Vincent, Payam Adineh, David Corney, Benno Stein, and Martin Potthast. 2019. SemEval-2019 Task 4: Hyperpartisan News Detection. In *Proceedings of the 13th International Workshop on Semantic Evaluation*, Jonathan May, Ekaterina Shutova, Aurelie Herbelot, Xiaodan Zhu, Marianna Apidianaki, and Saif M. Mohammad (Eds.). Association for Computational Linguistics, Minneapolis, Minnesota, USA, 829–839. https://doi.org/10.18653/v1/S19-2145

[20] Michelle YoungJin Kim and Kristen Marie Johnson. 2022. CLoSE: Contrastive Learning of Subframe Embeddings for Political Bias Classification of News Media. In *Proceedings of the 29th International Conference on Computational Linguistics*, Nicoletta Calzolari, Chu-Ren Huang, Hansaem Kim, James Pustejovsky, Leo Wanner, Key-Sun Choi, Pum-Mo Ryu, Hsin-Hsi Chen, Lucia Donatelli, Heng Ji, Sadao Kurohashi, Patrizia Paggio, Nianwen Xue, Seokhwan Kim, Younggyun Hahm, Zhong He, Tony Kyungil Lee, Enrico Santus, Francis Bond, and Seung-Hoon Na (Eds.). International Committee on Computational Linguistics, Gyeongju, Republic of Korea, 2780–2793. https://aclanthology.org/2022.coling-1.245

[21] Thomas Kipf and Max Welling. 2016. Variational Graph Auto-Encoders. *ArXiv* abs/1611.07308 (2016). https://api.semanticscholar.org/CorpusID:14249137

[22] Jan-David Krieger, Timo Spinde, Terry Ruas, Juhi Kulshrestha, and Bela Gipp. 2022. A domain-adaptive pre-training approach for language bias detection in news. In *Proceedings of the 22nd ACM/IEEE Joint Conference on Digital Libraries* (Cologne, Germany) *(JCDL '22)*. Association for Computing Machinery, New York, NY, USA, Article 3, 7 pages. https://doi.org/10.1145/3529372.3530932

[23] Quoc Le and Tomas Mikolov. 2014. Distributed Representations of Sentences and Documents. In *Proceedings of the 31st International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 32)*, Eric P. Xing and Tony Jebara (Eds.). PMLR, Bejing, China, 1188–1196. https://proceedings.mlr.press/v32/le14.html

[24] Yuanyuan Lei, Ruihong Huang, Lu Wang, and Nick Beauchamp. 2022. Sentence-level Media Bias Analysis Informed by Discourse Structures. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang (Eds.). Association for Computational Linguistics, Abu Dhabi, United Arab Emirates, 10040–10050. https://doi.org/10.18653/v1/2022.emnlp-main.682

[25] Luyang Lin, Lingzhi Wang, Xiaoyan Zhao, Jing Li, and Kam-Fai Wong. 2024. IndiVec: An Exploration of Leveraging Large Language Models for Media Bias Detection with Fine-Grained Bias Indicators. In *Findings of the Association for Computational Linguistics: EACL 2024*, Yvette Graham and Matthew Purver (Eds.). Association for Computational Linguistics, St. Julian's, Malta, 1038–1050. https://aclanthology.org/2024.findings-eacl.70

[26] Yonghao Liu, Renchu Guan, Fausto Giunchiglia, Yanchun Liang, and Xiaoyue Feng. 2021. Deep Attention Diffusion Graph Neural Networks for Text Classification. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih (Eds.). Association for Computational Linguistics, Online and Punta Cana, Dominican Republic, 8142–8152. https://doi.org/10.18653/v1/2021.emnlp-main.642

[27] Yujian Liu, Xinliang Frederick Zhang, David Wegsman, Nicholas Beauchamp, and Lu Wang. 2022. POLITICS: Pretraining with Same-story Article Comparison for Ideology Prediction and Stance Detection. In *Findings of the Association for Computational Linguistics: NAACL 2022*, Marine Carpuat, Marie-Catherine de Marneffe, and Ivan Vladimir Meza Ruiz (Eds.). Association for Computational Linguistics, Seattle, United States, 1354–1374. https://doi.org/10.18653/v1/2022.findings-naacl.101

[28] Giannis Nikolentzos, Antoine Tixier, and Michalis Vazirgiannis. 2020. Message Passing Attention Networks for Document Understanding. *Proceedings of the AAAI Conference on Artificial Intelligence* 34, 05 (April 2020), 8544–8551. https://doi.org/10.1609/aaai.v34i05.6376

[29] OpenAI. 2021. GPT-3.5 Turbo. (2021). https://platform.openai.com/docs/models/gpt-3-5-turbo unpublished, https://platform.openai.com/docs/models/gpt-3-5-turbo.

[30] Yinhua Piao, Sangseon Lee, Dohoon Lee, and Sun Kim. 2022. Sparse Structure Learning via Graph Neural Networks for Inductive Document Classification. *Proceedings of the AAAI Conference on Artificial Intelligence* 36, 10 (June 2022),

11165–11173. https://doi.org/10.1609/aaai.v36i10.21366

[31] Hannah Rabinowitz. 2016. Oath Keepers lawyer and federal judge in screaming match in court over US Capitol riot case. *CNN* (2016). https://www.cnn.com/2022/09/14/politics/oath-keepers-judge-mehta-argument/index.html

[32] Shaina Raza, Deepak John Reji, and Chen Ding. 2022. Dbias: detecting biases and ensuring fairness in news articles. *International Journal of Data Science and Analytics* 17, 1 (Sept. 2022), 39–59. https://doi.org/10.1007/s41060-022-00359-4

[33] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Kentaro Inui, Jing Jiang, Vincent Ng, and Xiaojun Wan (Eds.). Association for Computational Linguistics, Hong Kong, China, 3982–3992. https://doi.org/10.18653/v1/D19-1410

[34] Alan Roappeport. 2016. Vicente Fox and Felipe Calderón, 2 Ex-Presidents of Mexico, Say No Way Country Is Paying for Donald Trump's Wall. *NY Times* (2016). https://archive.nytimes.com/www.nytimes.com/politics/first-draft/2016/02/25/2-ex-presidents-of-mexico-say-no-way-theyre-paying-for-donald-trumps-wall/

[35] Francisco-Javier Rodrigo-Ginés, Jorge Carrillo de Albornoz, and Laura Plaza. 2023. A systematic review on media bias detection: What is media bias, how it is expressed, and how to detect it. *Expert Syst. Appl.* 237 (2023), 121641. https://api.semanticscholar.org/CorpusID:262174964

[36] Shamik Roy and Dan Goldwasser. 2020. Weakly Supervised Learning of Nuanced Frames for Analyzing Polarization in News Media. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Bonnie Webber, Trevor Cohn, Yulan He, and Yang Liu (Eds.). Association for Computational Linguistics, Online, 7698–7716. https://doi.org/10.18653/v1/2020.emnlp-main.620

[37] Benedek Rozemberczki, Oliver Kiss, and Rik Sarkar. 2020. Karate Club: An API Oriented Open-Source Python Framework for Unsupervised Learning on Graphs. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (Virtual Event, Ireland) *(CIKM '20)*. Association for Computing Machinery, New York, NY, USA, 3125–3132. https://doi.org/10.1145/3340531.3412757

[38] Timo Spinde, Christin Jeggle, Magdalena Haupt, Wolfgang Gaissmaier, and Helge Giese. 2022. How do we raise media bias awareness effectively? Effects of visualizations to communicate bias. *PLOS ONE* 17, 4 (04 2022), 1–14. https://doi.org/10.1371/journal.pone.0266204

[39] Timo Spinde, Jan-David Krieger, Terry Ruas, Jelena Mitrović, Franz Götz-Hahn, Akiko Aizawa, and Bela Gipp. 2022. Exploiting Transformer-Based Multitask Learning for the Detection of Media Bias in News Articles. In *Information for a Better World: Shaping the Global Future*, Malte Smits (Ed.). Springer International Publishing, Cham, 225–235.

[40] Timo Spinde, Manuel Plank, Jan-David Krieger, Terry Ruas, Bela Gipp, and Akiko Aizawa. 2021. Neural Media Bias Detection Using Distant Supervision With BABE - Bias Annotations By Experts. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih (Eds.). Association for Computational Linguistics, Punta Cana, Dominican Republic, 1166–1177. https://doi.org/10.18653/v1/2021.findings-emnlp.101

[41] Timo Spinde, Lada Rudnitckaia, Jelena Mitrović, Felix Hamborg, Michael Granitzer, Bela Gipp, and Karsten Donnay. 2021. Automated identification of bias inducing words in news articles using linguistic and context-oriented features. *Information Processing & Management* 58, 3 (may 2021), 102505. https://doi.org/10.1016/j.ipm.2021.102505

[42] Ciprian-Octavian Truică and Elena-Simona Apostol. 2023. It's All in the Embedding! Fake News Detection Using Document Embeddings. *Mathematics* 11, 3 (Jan. 2023), 508. https://doi.org/10.3390/math11030508

[43] Gisela Vallejo, Timothy Baldwin, and Lea Frermann. 2024. Connecting the Dots in News Analysis: Bridging the Cross-Disciplinary Disparities in Media Bias and Framing. In *Proceedings of the Sixth Workshop on Natural Language Processing and Computational Social Science (NLP+CSS 2024)*, Dallas Card, Anjalie Field, Dirk Hovy, and Katherine Keith (Eds.). Association for Computational Linguistics,

Mexico City, Mexico, 16–31. https://doi.org/10.18653/v1/2024.nlpcss-1.2

[44] Francielle Vargas, Kokil Jaidka, Thiago Pardo, and Fabrício Benevenuto. 2023. Predicting Sentence-Level Factuality of News and Bias of Media Outlets. In *Proceedings of the 14th International Conference on Recent Advances in Natural Language Processing*, Ruslan Mitkov and Galia Angelova (Eds.). INCOMA Ltd., Shoumen, Bulgaria, Varna, Bulgaria, 1197–1206. https://aclanthology.org/2023.ranlp-1.127

[45] Kunze Wang, Yihao Ding, and Soyeon Caren Han. 2024. Graph Neural Networks for Text Classification: A Survey. *Artificial Intelligence Review* 57, 190 (2024). https://link.springer.com/article/10.1007/s10462-024-10808-0#citeas

[46] K. Wang, S. Han, and J. Poon. 2022. InducT-GCN: Inductive Graph Convolutional Networks for Text Classification. In *2022 26th International Conference on Pattern Recognition (ICPR)*. IEEE Computer Society, Los Alamitos, CA, USA, 1243–1249. https://doi.org/10.1109/ICPR56361.2022.9956075

[47] Martin Wessel, Tomás Horych, Terry Ruas, Akiko Aizawa, Bela Gipp, and Timo Spinde. 2023. Introducing MBIB - The First Media Bias Identification Benchmark Task and Dataset Collection. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Taipei, Taiwan) *(SIGIR '23)*. Association for Computing Machinery, New York, NY, USA, 2765–2774. https://doi.org/10.1145/3539618.3591882

[48] Liang Yao, Chengsheng Mao, and Yuan Luo. 2019. Graph convolutional networks for text classification. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence* (Honolulu, Hawaii, USA) *(AAAI'19/IAAI'19/EAAI'19)*. AAAI Press, Article 905, 8 pages. https://doi.org/10.1609/aaai.v33i01.33017370

[49] Yawei Zhao and Xiaoyang Song. 2023. TextGCL: Graph Contrastive Learning for Transductive Text Classification. In *2023 International Joint Conference on Neural Networks (IJCNN)*. 1–8. https://doi.org/10.1109/IJCNN54540.2023.10191923

[50] Kaijian Zou, Xinliang Zhang, Winston Wu, Nicholas Beauchamp, and Lu Wang. 2023. Crossing the Aisle: Unveiling Partisan and Counter-Partisan Events in News Reporting. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, Houda Bouamor, Juan Pino, and Kalika Bali (Eds.). Association for Computational Linguistics, Singapore, 621–632. https://doi.org/10.18653/v1/2023.findings-emnlp.45

Disclaimer

The views expressed in this paper are those of the authors and do not reflect the official policy or position of the United States Army, the Department of Defense, or the United States Government.

Ethics Statement

All articles collected for this study were done so under the provisions of Section 107 of the U.S. Copyright Act and ensured that our collection action fell under the fair use category. The Tweets were collected in accordance with Twitter's terms of service at the time of collection. The code and news articles are made available at https://github.com/jhzsquared/DocNet__CodeData under the MIT and CC0 1.0 Universal licenses, respectively.

Acknowledgments