Multi-Task Ordinal Regression for Jointly Predicting the Trustworthiness and the Leading Political Ideology of News Media

Ramy Baly[1], **Georgi Karadzhov**[2], **Abdelrhman Saleh**[3], **James Glass**[1], **Preslav Nakov**[4]
[1]MIT Computer Science and Artificial Intelligence Laboratory, MA, USA
[2]SiteGround Hosting EOOD, Bulgaria, [3]Harvard University, MA, USA
[4]Qatar Computing Research Institute, HBKU, Qatar
{baly, glass}@mit.edu, georgi.m.karadjov@gmail.com
abdelrhman_saleh@college.harvard.edu, pnakov@hbku.edu.qa

Abstract

In the context of fake news, bias, and propaganda, we study two important but relatively under-explored problems: (*i*) *trustworthiness estimation* (on a 3-point scale) and (*ii*) *political ideology detection* (left/right bias on a 7-point scale) of entire news outlets, as opposed to evaluating individual articles. In particular, we propose a multi-task ordinal regression framework that models the two problems jointly. This is motivated by the observation that hyper-partisanship is often linked to low trustworthiness, e.g., appealing to emotions rather than sticking to the facts, while center media tend to be generally more impartial and trustworthy. We further use several auxiliary tasks, modeling centrality, hyper-partisanship, as well as left-vs.-right bias on a coarse-grained scale. The evaluation results show sizable performance gains by the joint models over models that target the problems in isolation.

1 Introduction

Recent years have seen the rise of social media, which has enabled people to virtually share information with a large number of users without regulation or quality control. On the bright side, this has given an opportunity for anyone to become a content creator, and has also enabled a much faster information dissemination. However, it has also opened the door for malicious users to spread disinformation and misinformation much faster, enabling them to easily reach audience at a scale that was never possible before. In some cases, this involved building sophisticated profiles for individuals based on a combination of psychological characteristics, meta-data, demographics, and location, and then micro-targeting them with personalized "fake news" with the aim of achieving some political or financial gains (Lazer et al., 2018; Vosoughi et al., 2018).

A number of fact-checking initiatives have been launched so far, both manual and automatic, but the whole enterprise remains in a state of crisis: by the time a claim is finally fact-checked, it could have reached millions of users, and the harm caused could hardly be undone. An arguably more promising direction is to focus on fact-checking entire news outlets, which can be done in advance. Then, we could fact-check the news before they were even written: by checking how trustworthy the outlets that published them are. Knowing the reliability of a medium is important not only when fact-checking a claim (Popat et al., 2017; Nguyen et al., 2018), but also when solving article-level tasks such as "fake news" and click-bait detection (Brill, 2001; Finberg et al., 2002; Hardalov et al., 2016; Karadzhov et al., 2017; De Sarkar et al., 2018; Pan et al., 2018; Pérez-Rosas et al., 2018)

Political ideology (or left/right bias) is a related characteristic, e.g., extreme left/right media tend to be propagandistic, while center media are more factual, and thus generally more trustworthy. This connection can be clearly seen in Figure 1.

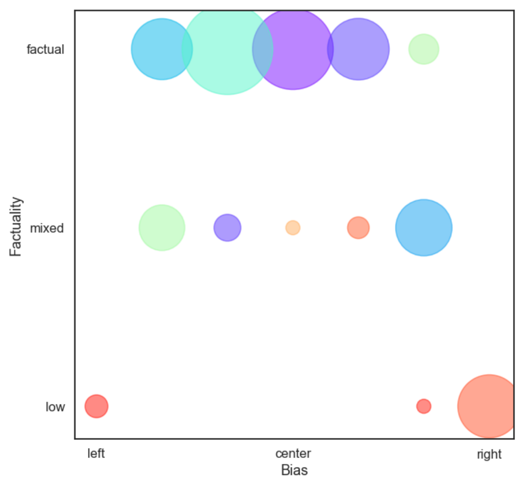

Figure 1: Correlation between bias and factuality for the news outlets in the Media Bias/Fact Check website.

Proceedings of NAACL-HLT 2019, pages 2109–2116
Minneapolis, Minnesota, June 2 - June 7, 2019.

Despite the connection between factuality and bias, previous research has addressed them as independent tasks, even when the underlying dataset had annotations for both (Baly et al., 2018). In contrast, here we solve them jointly. Our contributions can be summarized as follows:

- We study an under-explored but arguably important problem: predicting the factuality of reporting of news media. Moreover, unlike previous work, we do this jointly with the task of predicting political bias.
- As factuality and bias are naturally defined on an ordinal scale (factuality: from *low* to *high*, and bias: from *extreme-left* to *extreme-right*), we address them as ordinal regression. Using multi-task ordinal regression is novel for these tasks, and it is also an under-explored direction in machine learning in general.
- We design a variety of auxiliary subtasks from the bias labels: modeling centrality, hyper-partisanship, as well as left-vs.-right bias on a coarse-grained scale.

2 Related Work

Factuality of Reporting Previous work has modeled the factuality of reporting at the medium level by checking the general stance of the target medium with respect to known manually fact-checked claims, without access to gold labels about the overall medium-level factuality of reporting (Mukherjee and Weikum, 2015; Popat et al., 2016, 2017, 2018).

The trustworthiness of Web sources has also been studied from a Data Analytics perspective, e.g., Dong et al. (2015) proposed that a trustworthy source is one that contains very few false claims. In social media, there has been research targeting the user, e.g., finding malicious users (Mihaylov and Nakov, 2016; Mihaylova et al., 2018; Mihaylov et al., 2018), *sockpuppets* (Maity et al., 2017), *Internet water army* (Chen et al., 2013), and *seminar users* (Darwish et al., 2017).

Unlike the above work, here we study source reliability as a task in its own right, using manual gold annotations specific for the task and assigned by independent fact-checking journalists. Moreover, we address the problem as one of ordinal regression on a three-point scale, and we solve it jointly with political ideology prediction in a multi-task learning setup, using several auxiliary tasks.

Predicting Political Ideology In previous work, political ideology, also known as media bias, was used as a feature for "fake news" detection (Horne et al., 2018a). It has also been the target of classification, e.g., Horne et al. (2018b) predicted whether an article is biased (*political* or *bias*) vs. unbiased. Similarly, Potthast et al. (2018) classified the bias in a target article as (*i*) left vs. right vs. mainstream, or as (*ii*) hyper-partisan vs. mainstream. Left-vs-right bias classification at the article level was also explored by Kulkarni et al. (2018), who modeled both the textual and the URL contents of the target article. There has been also work targeting bias at the phrase or the sentence level (Iyyer et al., 2014), focusing on political speeches (Sim et al., 2013) or legislative documents (Gerrish and Blei, 2011), or targeting users in Twitter (Preoţiuc-Pietro et al., 2017). Another line of related work focuses on propaganda, which can be seen as a form of extreme bias (Rashkin et al., 2017; Barrón-Cedeño et al., 2019a,b). See also a recent position paper (Pitoura et al., 2018) and an overview paper on bias on the Web (Baeza-Yates, 2018). Unlike the above work, here we focus on predicting the political ideology of news media outlets.

In our previous work (Baly et al., 2018), we did target the political bias of entire news outlets, as opposed to working at the article level (we also modeled factuality of reporting, but as a separate task without trying multi-task learning). In addition to the text of the articles published by the target news medium, we used features extracted from its corresponding Wikipedia page and Twitter profile, as well as analysis of its URL structure and traffic information about it from Alexa rank. In the present work, we use a similar set of features, but we treat the problem as one of ordinal regression. Moreover, we model the political ideology and the factuality of reporting jointly in a multi-task learning setup, using several auxiliary tasks.

Multitask Ordinal Regression *Ordinal regression* is well-studied and is commonly used for text classification on an ordinal scale, e.g., for sentiment analysis on a 5-point scale (He et al., 2016; Rosenthal et al., 2017a). However, *multi-task ordinal regression* remains an understudied problem.

Yu et al. (2006) proposed a Bayesian framework for collaborative ordinal regression, and demonstrated that modeling multiple ordinal regression tasks outperforms single-task models.

Walecki et al. (2016) were interested in jointly predicting facial action units and their intensity level. They argued that, due to the high number of classes, modeling these tasks independently would be inefficient. Thus, they proposed the *copula ordinal regression* model for multi-task learning and demonstrated that it can outperform various single-task setups. We use this model in our experiments below.

Balikas et al. (2017) used multi-task ordinal regression for the task of fine-grained sentiment analysis. In particular, they introduced an auxiliary coarse-grained task on a 3-point scale, and demonstrated that it can improve the results for sentiment analysis on the original 5-point scale. Inspired by this, below we experiment with different granularity for political bias; however, we explore a larger space of possible auxiliary tasks.

3 Method

Copula Ordinal Regression We use the *Copula Ordinal Regression* (COR) model, which was originally proposed by Walecki et al. (2016) to estimate the intensities of facial action units (AUs). The model uses copula functions and conditional random fields (CRFs) to approximates the learning of the joint probability distribution function (PDF) of the facial AUs (random variables), using the bivariate joint distributions capturing dependencies between AU pairs. It was motivated by the fact that (*i*) many facial AUs co-exist with different levels of intensity, (*ii*) some AUs co-occur more often than others, and (*iii*) some AUs depend on the intensity of other units.

We can draw an analogy between modeling facial AUs and modeling news media, where each medium expresses a particular bias (political ideology) and can also be associated with a particular level of factuality. Therefore, bias and factuality can be analogous to the facial AUs in (Walecki et al., 2016), and represent two aspects of news reporting, each being modeled on a multi-point ordinal scale. In particular, we model bias on a 7-point scale (*extreme-left*, *left*, *center-left*, *center*, *center-right*, *right*, and *extreme-right*), and factuality on a 3-point scale (*low*, *mixed*, and *high*).

In our case, we train the COR model to predict the joint PDF between political bias and factuality of reporting. This could potentially work well given the inherent inter-dependency between the two tasks as we have seen on Figure 1.

Auxiliary Tasks We use a variety of auxiliary tasks, derived from the bias labels. This includes converting the 7-point scale to (*i*) 5-point and 3-point scales, similarly to (Balikas et al., 2017), and to (*ii*) a 2-point scale in two ways to model extreme partisanship, and centrality. Here is the list of the auxiliary tasks we use with precise definition of the label mappings:

- **Bias5-way:** Predict bias on a 5-pt scale; 1:*extreme-left*, 2:*left*, 3:{*center-left, center, center-right*}, 4:*right*, and 5:*extreme-right*.
- **Bias3-way:** Predict bias on a 3-pt scale; 1:{*extreme-left, left*}, 2:{*center-left, center, center-right*}, and 3:{*right, extreme-right*}.
- **Bias-extreme:** Predict extreme vs. non-extreme partisanship on a 2-pt scale; 1:{*extreme-left, extreme-right*}, 2:{*left, center-left, center, center-right, right*}.
- **Bias-center:** Predict center vs. non-center political ideology on a 2-pt scale, ignoring polarity: 1:{*extreme-left, left, right, extreme-right*}, 2:{*center-left, center, center-right*}.

Features We used the features from (Baly et al., 2018)[1]. We gathered a sample of articles from the target medium, and we calculated features such as POS tags, linguistic cues, sentiment scores, complexity, morality, as well as embeddings. We also used the Wikipedia page of the medium (if any) to generate document embedding. Then, we collected metadata from the medium's Twitter account (if any), e.g., whether is is verified, number of followers, whether the URL in the Twitter page matches the one of the medium. Finally, we added Web-based features that (*i*) model the orthographic structure of the medium's URL address, and (*ii*) analyze the Web-traffic information about the medium's website, as found in Alexa rank.[2]

4 Experiments and Evaluation

Data We used the MBFC dataset (Baly et al., 2018) that has 1,066 news media manually annotated for factuality (3-pt scale: *high*, *mixed*, *low*) and political bias (7-pt scale: from *extreme-left* to *extreme-right*). This dataset was annotated by volunteers using a detailed methodology[3] that is designed to guarantee annotation objectivity.

[1] https://github.com/ramybaly/News-Media-Reliability

[2] https://www.alexa.com/siteinfo

[3] For details, see https://mediabiasfactcheck.com/methodology/

Name	URL	Bias	Factuality	Twitter Handle	Wikipedia page
London Web News	londonwebnews.com	Extreme Left	Low	@londonwebnews	N/A
Daily Mirror	www.mirror.co.uk	Left	Mixed	@DailyMirror	~/Daily_Mirror
NBC News	www.nbcnews.com	Center-Left	High	@nbcnews	~/NBC_News
Associated Press	apnews.com	Center	Very High	@apnews	~/Associated_Press
Gulf News	gulfnews.com	Center-Right	High	@gulf_news	~/Gulf_News
Russia Insider	russia-insider.com	Right	Mixed	@russiainsider	~/Russia_Insider
Breitbart	www.breitbart.com	Extreme Right	Low	@BreitbartNews	~/Breitbart_News

Table 1: Examples of media and their labels for bias and factuality of reporting derived from MBFC.

Furthermore, readers can provide their own feedback on existing annotations, and in case of a large discrepancy, annotation is adjusted after a thorough review. Therefore, we believe the annotation quality is good enough to experiment with. We noticed that 117 media had *low* factuality because they publish *satire* and *pseudo-science*, neither of which has a political perspective. Since we are interested in modeling the relation between factuality and bias, we excluded those websites, thus ending up with 949 news media. Some examples from this dataset are shown in Table 1 with both factuality and bias labels, in addition to their corresponding Twitter handles and Wikipedia pages. Overall, 64% of the media in our dataset have Wikipedia pages, and 65% have Twitter accounts. Table 2 further provides detailed statistics about the label distribution in the MBFC dataset.

Factuality		Bias	
Low	198	Extreme-Left	23
Mixed	282	Left	151
High	469	Center-Left	200
		Center	139
		Center-Right	105
		Right	164
		Extreme-Right	167

Table 2: Label distribution (counts) in the MBFC dataset, which we used in our experiments.

Experimental Setup We used the implementation[4] of the Copula Ordinal Regression (COR) model as described in (Walecki et al., 2016). In our experiments, we used 5-fold cross-validation, where for each fold we split the training dataset into a training part and a validation part, and we used the latter to fine-tune the model's hyper-parameters, optimizing for Mean Absolute Error (MAE). MAE is an appropriate evaluation measure given the ordinal nature of the tasks.

These hyper-parameters include the copula function (*Gumbel* vs. *Frank*), the marginal distribution (*normal* vs. *sigmoid*), the number of training iterations, the optimizer (*gradient descent*, *BFGS*), and the connection density of the CRFs. We report both MAE and MAE^M, which is a variant of MAE that is more robust to class imbalance. See (Baccianella et al., 2009; Rosenthal et al., 2017b) for more details about MAE^M vs. MAE. We compare the results to two baselines: (*i*) majority class, and (*ii*) single-task ordinal regression.

Results and Discussion Table 3 shows the evaluation results for the COR model when trained to jointly model the main task (*shown in the columns*) using combinations of auxiliary tasks (*shown in the rows*). We can see that the single-task ordinal regression model performs much better than the majority class baseline based on both evaluation measures. We can further see that the performance on the main task improves when jointly modeling several auxiliary tasks. This improvement depends on the auxiliary tasks in use.

For factuality prediction, it turns out that the combination of *bias-center*+*bias-extreme* yields the best overall MAE of 0.481. This makes sense and aligns well with the intuition that knowing whether a medium is centric or hyper-partisan is important to predict the factuality of its reporting. For instance, a news medium without a political ideology tends to be more trustworthy compared to an extremely biased one, regardless of their polarity (left or right), as we should expect based on the data distribution shown in Figure 1 above.

For bias prediction (at a 7-point left-to-right scale), a joint model that uses political bias at different levels of granularity (5-point and 3-point) as auxiliary tasks yields the best overall MAE of 1.479. This means that jointly modeling bias with the same information at coarser levels of granularity, i.e., adding 3-point and 5-point as auxiliary tasks, reduces the number of gross mistakes.

[4] https://github.com/RWalecki/copula_ordinal_regression

	Factuality		Bias	
Auxiliary Tasks	**MAE**	**MAE**^{M}	**MAE**	**MAE**^{M}
(None) majority class	0.714	1.000	1.798	1.857
(None) single-task COR	0.514	0.567	1.582	1.728
+bias	0.526	0.566	–	–
+factuality	–	–	1.584	1.695
+bias5-way	0.495	0.541	1.504 *(1.485)*	1.627 *(1.647)*
+bias3-way	0.497	0.548	1.528 *(1.498)*	1.658 *(1.654)*
+bias-center	0.509	0.561	1.594 *(1.535)*	1.745 *(1.695)*
+bias-extreme	0.498	0.550	1.584 *(1.558)*	1.743 *(1.726)*
+bias5-way+bias3-way	0.493	0.541	1.479 ***(1.475)***	1.637 ***(1.623)***
+bias-center+bias-extreme	**0.481**	**0.529**	1.563 *(1.526)*	1.714 *(1.672)*
+bias5-way+bias3-way+bias-center+bias-extreme	0.485	0.537	1.513 *(1.504)*	1.665 *(1.677)*

Table 3: Evaluating the copula ordinal regression model trained to jointly model the main task (*shown in the columns*) and different auxiliary tasks (*shown in the rows*). The results in parentheses correspond to the case when factuality is added as an additional auxiliary task (only applicable when the main task is bias prediction).

E.g., predicting *extreme-left* instead of *extreme-right*, since the model is encouraged by the auxiliary tasks to learn the correct polarity, regardless of its intensity. We can see that *factuality* is not very useful as an auxiliary task by itself (MAE=1.584 and MAE^M=1.695). In other words, a medium with low factuality could be extremely biased to either the right or to the left. Therefore, relying on *factuality* alone to predict bias might introduce severe errors, e.g., confusing extreme-left with extreme-right, thus leading to higher MAE scores. This can be remedied by adding *factuality* to the mix of other auxiliary tasks to model the main task (7-point bias prediction). The results of these experiments, shown in parentheses in Table 3, indicate that adding *factuality* to any combination of auxiliary tasks consistently yields lower MAE scores. In particular, modeling the combination of *factuality+bias5-way+bias3-way* yields the best results (MAE=1.475 and MAE^M=1.623). This result indicates that *factuality* provides complementary information that can help predict bias.

We ran a two-tailed t-test for statistical significance, which is suitable for an evaluation measure such as MAE, to confirm the improvements that were introduced by the multi-task setup. We found that the best models (shown in bold in Table 3) outperformed both the corresponding majority class baselines with a p-value ≤ 0.001, and the corresponding single-task ordinal regression baselines with a p-value ≤ 0.02.

Finally, we compared the above results to our previous work (Baly et al., 2018) by independently training a Support Vector Machine (SVM) classifier for each task, using the same features. The resulting MAE was 0.450 for factuality and 1.184 for bias prediction, which is slightly better then our results (yet, very comparable for factuality). However, our goal here is to emphasize the advantages of modeling the two tasks jointly.

5 Conclusion and Future Work

We have presented a multi-task ordinal regression framework for jointly predicting trustworthiness and political ideology of news media sources, using several auxiliary tasks, e.g., based on a coarser-grained scales or modeling extreme partisanship. Overall, we have observed sizable performance gains in terms of reduced MAE by the multi-task ordinal regression models over single-task models for each of the two individual tasks.

In future work, we want to try more auxiliary tasks, and to experiment with other languages. We further plan to go beyond *left vs. right*, which is not universal and can exhibit regional specificity (Tavits and Letki, 2009), and to model other kinds of biases, e.g., *eurosceptic vs. europhile*, *nationalist* vs. *globalist*, *islamist vs. secular*, etc.

Acknowledgments

This research is part of the Tanbih project,[5] which aims to limit the effect of “fake news”, propaganda and media bias by making users aware of what they are reading. The project is developed in collaboration between the MIT Computer Science and Artificial Intelligence Laboratory (CSAIL) and the Qatar Computing Research Institute (QCRI), HBKU.

[5] http://tanbih.qcri.org/

References

Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2009. Evaluation measures for ordinal regression. In *Proceedings of the 9th IEEE International Conference on Intelligent Systems Design and Applications*, ISDA '09, pages 283–287, Pisa, Italy.

Ricardo Baeza-Yates. 2018. Bias on the web. *Commun. ACM*, 61(6):54–61.

Georgios Balikas, Simon Moura, and Massih-Reza Amini. 2017. Multitask learning for fine-grained Twitter sentiment analysis. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '17, pages 1005–1008, Tokyo, Japan.

Ramy Baly, Georgi Karadzhov, Dimitar Alexandrov, James Glass, and Preslav Nakov. 2018. Predicting factuality of reporting and bias of news media sources. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '18, pages 3528–3539, Brussels, Belgium.

Alberto Barrón-Cedeño, Giovanni Da San Martino, Israa Jaradat, and Preslav Nakov. 2019a. Proppy: A system to unmask propaganda in online news. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence*, AAAI'19, Honolulu, HI, USA.

Alberto Barrón-Cedeño, Giovanni Da San Martino, Israa Jaradat, and Preslav Nakov. 2019b. Proppy: Organizing news coverage on the basis of their propagandistic content. *Information Processing and Management*.

Ann M Brill. 2001. Online journalists embrace new marketing function. *Newspaper Research Journal*, 22(2):28.

Cheng Chen, Kui Wu, Venkatesh Srinivasan, and Xudong Zhang. 2013. Battling the Internet Water Army: detection of hidden paid posters. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, ASONAM '13, pages 116–120, Niagara, Canada.

Kareem Darwish, Dimitar Alexandrov, Preslav Nakov, and Yelena Mejova. 2017. Seminar users in the Arabic Twitter sphere. In *Proceedings of the 9th International Conference on Social Informatics*, SocInfo '17, pages 91–108, Oxford, UK.

Sohan De Sarkar, Fan Yang, and Arjun Mukherjee. 2018. Attending sentences to detect satirical fake news. In *Proceedings of the 27th International Conference on Computational Linguistics*, COLING '18, pages 3371–3380, Santa Fe, NM, USA.

Xin Luna Dong, Evgeniy Gabrilovich, Kevin Murphy, Van Dang, Wilko Horn, Camillo Lugaresi, Shaohua Sun, and Wei Zhang. 2015. Knowledge-based trust: Estimating the trustworthiness of web sources. *Proc. VLDB Endow.*, 8(9):938–949.

Howard Finberg, Martha L Stone, and Diane Lynch. 2002. Digital journalism credibility study. *Online News Association. Retrieved November*, 3:2003.

Sean M. Gerrish and David M. Blei. 2011. Predicting legislative roll calls from text. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ICML '11, pages 489–496, Bellevue, Washington, USA.

Momchil Hardalov, Ivan Koychev, and Preslav Nakov. 2016. In search of credible news. In *Proceedings of the 17th International Conference on Artificial Intelligence: Methodology, Systems, and Applications*, AIMSA '16, pages 172–180, Varna, Bulgaria.

Yunchao He, Liang-Chih Yu, Chin-Sheng Yang, K Robert Lai, and Weiyi Liu. 2016. YZU-NLP team at semeval-2016 task 4: Ordinal sentiment classification using a recurrent convolutional network. In *Proceedings of the 10th International Workshop on Semantic Evaluation*, SemEval '16, pages 251–255, San Diego, CA, USA.

Benjamin Horne, Sara Khedr, and Sibel Adali. 2018a. Sampling the news producers: A large news and feature data set for the study of the complex media landscape. In *Proceedings of the Twelfth International Conference on Web and Social Media*, ICWSM '18, pages 518–527, Stanford, CA, USA.

Benjamin D. Horne, William Dron, Sara Khedr, and Sibel Adali. 2018b. Assessing the news landscape: A multi-module toolkit for evaluating the credibility of news. In *Proceedings of the The Web Conference*, WWW '18, pages 235–238, Lyon, France.

Mohit Iyyer, Peter Enns, Jordan Boyd-Graber, and Philip Resnik. 2014. Political ideology detection using recursive neural networks. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pages 1113–1122, Baltimore, MD, USA.

Georgi Karadzhov, Pepa Gencheva, Preslav Nakov, and Ivan Koychev. 2017. We built a fake news & click-bait filter: What happened next will blow your mind! In *Proceedings of the International Conference on Recent Advances in Natural Language Processing*, RANLP '17, pages 334–343, Varna, Bulgaria.

Vivek Kulkarni, Junting Ye, Steven Skiena, and William Yang Wang. 2018. Multi-view models for political ideology detection of news articles. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '18, pages 3518–3527, Brussels, Belgium.

David M.J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, and Jonathan L. Zittrain. 2018. The science of fake news. *Science*, 359(6380):1094–1096.

Suman Kalyan Maity, Aishik Chakraborty, Pawan Goyal, and Animesh Mukherjee. 2017. Detection of sockpuppets in social media. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, pages 243–246, Portland, OR, USA.

Todor Mihaylov, Tsvetomila Mihaylova, Preslav Nakov, Lluís Màrquez, Georgi Georgiev, and Ivan Koychev. 2018. The dark side of news community forums: Opinion manipulation trolls. *Internet Research*, 28(5):1292–1312.

Todor Mihaylov and Preslav Nakov. 2016. Hunting for troll comments in news community forums. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, ACL '16, pages 399–405, Berlin, Germany.

Tsvetomila Mihaylova, Preslav Nakov, Lluís Màrquez, Alberto Barrón-Cedeño, Mitra Mohtarami, Georgi Karadjov, and James Glass. 2018. Fact checking in community forums. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, AAAI '18, pages 879–886, New Orleans, LA, USA.

Subhabrata Mukherjee and Gerhard Weikum. 2015. Leveraging joint interactions for credibility analysis in news communities. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, CIKM '15, pages 353–362, Melbourne, Australia.

An T. Nguyen, Aditya Kharosekar, Matthew Lease, and Byron C. Wallace. 2018. An interpretable joint graphical model for fact-checking from crowds. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, AAAI '18, New Orleans, LA, USA.

Jeff Z. Pan, Siyana Pavlova, Chenxi Li, Ningxi Li, Yangmei Li, and Jinshuo Liu. 2018. Content based fake news detection using knowledge graphs. In *Proceedings of the International Semantic Web Conference*, ISWC '18, Monterey, CA, USA.

Verónica Pérez-Rosas, Bennett Kleinberg, Alexandra Lefevre, and Rada Mihalcea. 2018. Automatic detection of fake news. In *Proceedings of the 27th International Conference on Computational Linguistics*, COLING '18, pages 3391–3401, Santa Fe, NM, USA.

Evaggelia Pitoura, Panayiotis Tsaparas, Giorgos Flouris, Irini Fundulaki, Panagiotis Papadakos, Serge Abiteboul, and Gerhard Weikum. 2018. On measuring bias in online information. *SIGMOD Rec.*, 46(4):16–21.

Kashyap Popat, Subhabrata Mukherjee, Jannik Strötgen, and Gerhard Weikum. 2016. Credibility assessment of textual claims on the web. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, CIKM '16, pages 2173–2178, Indianapolis, IN, USA.

Kashyap Popat, Subhabrata Mukherjee, Jannik Strötgen, and Gerhard Weikum. 2017. Where the truth lies: Explaining the credibility of emerging claims on the Web and social media. In *Proceedings of the 26th International Conference on World Wide Web Companion*, WWW '17, pages 1003–1012, Perth, Australia.

Kashyap Popat, Subhabrata Mukherjee, Jannik Strötgen, and Gerhard Weikum. 2018. CredEye: A credibility lens for analyzing and explaining misinformation. In *Proceedings of The Web Conference 2018*, WWW '18, pages 155–158, Lyon, France.

Martin Potthast, Johannes Kiesel, Kevin Reinartz, Janek Bevendorff, and Benno Stein. 2018. A stylometric inquiry into hyperpartisan and fake news. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, ACL '18, pages 231–240, Melbourne, Australia.

Daniel Preoţiuc-Pietro, Ye Liu, Daniel Hopkins, and Lyle Ungar. 2017. Beyond binary labels: Political ideology prediction of Twitter users. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, ACL '17, pages 729–740, Vancouver, Canada.

Hannah Rashkin, Eunsol Choi, Jin Yea Jang, Svitlana Volkova, and Yejin Choi. 2017. Truth of varying shades: Analyzing language in fake news and political fact-checking. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, EMNLP '17, pages 2931–2937, Copenhagen, Denmark.

Sara Rosenthal, Noura Farra, and Preslav Nakov. 2017a. SemEval-2017 task 4: Sentiment analysis in Twitter. In *Proceedings of the 11th International Workshop on Semantic Evaluation*, SemEval '17, pages 502–518, Vancouver, Canada.

Sara Rosenthal, Noura Farra, and Preslav Nakov. 2017b. SemEval-2017 task 4: Sentiment analysis in Twitter. In *Proceedings of the 11th International Workshop on Semantic Evaluation*, SemEval '17, pages 502–518, Vancouver, Canada.

Yanchuan Sim, Brice D. L. Acree, Justin H. Gross, and Noah A. Smith. 2013. Measuring ideological proportions in political speeches. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, EMNLP '13, pages 91–101, Seattle, WA, USA.

Margit Tavits and Natalia Letki. 2009. When left is right: Party ideology and policy in Post-Communist Europe. *The American Political Science Review*, 103(4):555–569.

Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science*, 359(6380):1146–1151.

Robert Walecki, Ognjen Rudovic, Vladimir Pavlovic, and Maja Pantic. 2016. Copula ordinal regression for joint estimation of facial action unit intensity. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4902–4910.

Shipeng Yu, Kai Yu, Volker Tresp, and Hans-Peter Kriegel. 2006. Collaborative ordinal regression. In *Proceedings of the 23rd international conference on Machine learning*, pages 1089–1096. ACM.