

Aplicação de Reconhecimento de Entidades Nomeadas para Detecção de *Fake News*

Marcos A. Spalenza

Programa de Pós-Graduação em Informática - PPGI

Universidade Federal do Espírito Santo - Vitória/ES - Brasil

Abstract

As *fake news* são problemas atuais de disseminação de conteúdo incoerente e fraudulento em meio da mídia tradicional e das redes sociais. Sabendo que a capacidade de produção é muito superior das agências de verificação dos fatos, torna-se necessário o apoio da detecção automática deste tipo de conteúdo. Portanto, neste artigo propomos uma abordagem de análise estrutural linguística com reconhecimento de entidades para identificação das *fake news*. Através dessa abordagem identificamos estruturas linguísticas que devem diferenciar um artigo produzido pelas agências de notícias formais da produção falsa. Nesse aspecto, apresentamos avaliações similares aos métodos com uso de informações externas, alcançando resultados médios de 90% de acurácia nos datasets da literatura.

Keywords: Fake News, Processamento de Linguagem Natural, Classificação de Documentos, Extração de Informação

1. Introdução

Os processos de análise de documentos e verificação de conteúdo há muito tempo são empregadas para controle, organização e busca documental. Essa análise documental há muito tempo é empregada para revisão textual, validação do formato e garantia de contextualização e coesão. Dentro desses métodos, a revisão textual continuamente vem sendo automatizada e os computadores se tornam proficientes para definir incoerências linguísticas na sequência textual. Então, a área de *Natural Language Processing - NLP* é demandada com mais ênfase para a verificação de conteúdo e garantia de coerência.

Dentre os desafios da análise escrita, a detecção de fatos incoerentes vêm ganhando crescente atenção. Essa atenção é devida ao impacto socio-político de vários países do mundo e, por consequência, ganha notoriedade legislativa nos mesmos [1]. No Brasil, um recente Inquérito das *Fake News* do Supremo Tribunal Federal - STF demonstra a relevância desse debate no âmbito popular e jurídico. Esse processo investiga possíveis ataques e *fake news* promovidos em sites e redes sociais com ataques ao STF. Em decorrência disso, o projeto de lei PL 2.630/2020¹ cria a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet, com as premissas de combate às *fake news* e a responsabilização pelo conteúdo. Assim, a automação do processo de identificação, filtragem e validação de conteúdos se torna notório e a credibilidade é cada vez mais requisitada.

Impulsionada por *bots*, as notícias veiculam dentro de redes sociais unindo-se com versões distorcidas, opinativas e tendenciosas. Portanto, além de notícias elaboradas com conteúdo

enganoso, notícias verdadeiras recebem interpretações falsas. Apesar da importância da credibilidade do conteúdo, a checagem manual dos dados não é possível em larga escala [2]. Para atender parcialmente essa demanda, as agências de verificação surgiram para validação das notícias com maior veiculação [3, 4]. Portanto, a automação do processo, a popularização da checagem e a garantia de credibilidade são esperadas pelo público enquanto a produção de conteúdo preza pela alcunha do seu crédito.

A diferença de uma notícia factual e uma notícia falsa têm suas nuances. A mudança de contexto, a fonte de origem, a descaracterização ou a omissão do fato são métodos comuns para lesar os citados [5]. Para além disso, é conhecida que tais notícias apresentam linguagem própria, frequentemente com erros ou incoerências e personagens específicos da sociedade. Dentre estes personagens, encontram-se políticos, empresas reconhecidas, instituições governamentais, locais e entidades com potencial jornalístico [?]. Apesar do problema das imagens e áudios forjados esse trabalho se propõe a apresentar uma ferramenta para checagem de fatos através da análise sintática. Com esse tipo de análise textual, esperamos reconhecer estruturas na sintaxe e no método de escrita que caracterizem o formato das *fake news*, para além do seu conteúdo.

Atualmente investigamos o estilo e formato de escrita de textos incoerentes e enganosos. Essa abordagem relaciona-se com a verificação da qualidade textual. De qualquer forma, a escrita enganosa é aqui analisada de forma isentada de características sociais, informações externas ou politização. O texto, portanto, é tratado pelos métodos de *Part-of-Speech Tagging - POS-Tag* e *Named Entity Recognition - NER*.

O artigo descreve o modelo proposto em 6 seções. Na Seção 2, descrevemos trabalhos correlatos e métodos utilizados na literatura para a classificação de *fake news*. Na Seção 3, propo-

¹ Email address: marcos.spalenza@gmail.com (Marcos A. Spalenza)

¹ Projeto de Lei 2.630/2020 <https://www.camara.leg.br/proposicoes-web/fichadetramitacao?idProposicao=2256735>

mos o modelo de interpretação sintática dos documentos. Na Seção 4, caracterizamos os datasets que foram utilizados pelos experimentos em português e inglês. Na Seção 5 apresentamos os resultados obtidos com cada dataset. Por fim, na Seção 6, elencamos nossas conclusões e algumas perspectivas de trabalhos futuros.

2. Trabalhos Relacionados

O processo de identificação das *fake news* inclui a combinação de portais de notícias e a disseminação do conteúdo nas mídias sociais [6]. É relevante também destacar a importância da origem, mídia e interações sociais que relacionam-se com a notícia [7].

Portanto, em redes sociais métodos distintos de investigação vão além da análise do conteúdo. De forma temporal e relacional, é verificada a relação entre usuários e as interações que tornam a notícia relevante dentro da rede social [8]. Enquanto, outras técnicas tentam associar as notícias originais à proliferação do conteúdo na rede. Esses métodos observam o conteúdo original e avaliam a repercussão nas demais mídias [9].

Por conta disso, para determinadas abordagens a análise dos metadados pode significar aumento relevante da qualidade da classificação. Assim, para considerar fontes relevantes, sejam usuários acreditados [9] ou agências de notícia relevantes [10, 11], relaciona-se texto como uma assinatura da escrita em relação à fonte.

Apesar de tudo, os modelos linguísticos são os mais comuns por sua independência de tempo e interações em mídia. Esses tipos consideram a divulgação do conteúdo já disseminado em meio a mídia e o alcance em larga escala. Abordagens textuais mais comuns analisam a relação entre os termos e a veracidade dos fatos [12]. Porém, estratégias com acréscimo de estruturas relacionais entre os termos e funções léxico-sintáticas na criação do texto evidenciaram-se.

As nuances deste tipo de escrita, caracterizam o autor e assinam a *fake news*. Aspectos descritivos destacam-se como humor, complexidade da sentença, pontuação e negação [13]. Através disso, os aspectos estruturais da linguagem tornam métodos inter-domínio [14] e inter-linguagem [15] hábeis identificadores de possíveis conteúdos inverídicos. Nessa perspectiva, estes identificadores da escrita descrevem assinaturas, não só do autor, quanto do seu objetivo e seu público-alvo.

3. O Modelo

A leitura da informação é feita de forma a interpretar a função sintática de cada palavra do documento com *POS-Tag*. Ainda com o texto, o processo de *NER* identifica entidades, em geral representações dos nomes encontrados nos documentos. Os nomes são incluídos na posição do rótulo sintático da palavra, como uma especificação dos substantivos - NOUN e substantivos próprios - PNOUN. Para exemplo deste processamento o documento de exemplo é dado a seguir:

- Maior cientista do mundo alerta: "O Planeta Terra estará inabitável no ano de 3016". Stephen Hawking, em entrevista, alertou mais uma vez que a humanidade corre perigo. A tecnologia e os domínios da ciência são os fatores cruciais para que o planeta Terra seja extinto. ²

- Planeta Terra - LOC, Stephen Hawking - PER, Terra - LOC

- Maior ADJ cientista NOUN do DET mundo NOUN alerta ADJ : PUNCT "PUNCT O DET Planeta PROPN Terra PROPN estará VERB inabitável ADJ no DET ano NOUN de ADP 3016 NUM "NOUN . PUNCT SPACE Stephen PROPN Hawking PROPN , PUNCT SPACE em ADP entrevista NOUN , PUNCT SPACE alertou VERB mais ADV uma ADP vez NOUN que CONJ a DET humanidade NOUN corre VERB perigo NOUN . PUNCT A DET tecnologia NOUN e CONJ os DET domínios SYM da ADP ciência NOUN são VERB os DET fatores SYM cruciais ADJ para ADP que CONJ o DET planeta NOUN Terra PROPN seja VERB extinto VERB . PUNCT

- ADJ NOUN DET NOUN ADJ PUNCT PUNCT DET **LOC LOC** VERB ADJ DET NOUN ADP NUM NOUN PUNCT SPACE **PER PER** PUNCT SPACE ADP NOUN PUNCT SPACE VERB ADV ADP NOUN CONJ DET NOUN VERB NOUN PUNCT DET NOUN CONJ DET SYM ADP NOUN VERB DET SYM ADJ ADP CONJ DET NOUN **LOC** VERB VERB PUNCT

O exemplo mostra as três fazes do documento, a entrada textual, a estrutura sintática dos termos e, por fim, o reconhecimento dos nomes. Com este formato é aplicada a análise de frequência por janelas de 1 a 5 termos, determinadas *n-grams*. Pelo padrão dos *n-grams* então, esperamos encontrar estruturas frasais clássicas das *fake news*, características que mostrem associações entre entidades e possíveis erros na estrutura deste tipo de texto. Vale ainda ressaltar que, associado à isso, nós não utilizamos nesta abordagem a frequência dos termos, os metadados com fontes e origens dos documentos ou informações externas aos documentos. Estes demais conteúdos podem representar vieses para a classificação [6]. Exemplos de viés da classificação com inferência textual ou por metadado são os limites de diversidade do domínio, de tamanho do texto e vocabulário ou a tendência e frequência das fontes [10].

Associado a isso, o *NER* caracteriza-se por ser um meio de acrescentar estrutura semântica para análise de documentos. Este método indica entidades relevantes que, semanticamente, podem ser interpretadas computacionalmente como associações mais fortes entre os termos. Assim, bem como apresentado no documento de exemplo, entidades ressaltam e são indicados na classificação com marcações específicas. As marcações feitas para ambas as linguagens indicaram Pessoas - PER, Organizações - ORG, Locais - LOC e Variadas - MISC. Nas anotações MISC encontram-se outras classes juntas em um mesmo grupo como tempo, valor, obra de arte, acontecimento dentre outros.

²Extrato de um dos documentos do dataset *Fake.BR*

165 Enquanto as entidades nomeadas indicam alvos em potencial da notícia, as marcações sintáticas caracterizam como é feita a construção factual e a relação entre as entidades. A análise conjunta de *POSTag+NER* visa acrescentar contexto para as entidades nomeadas, dando maior destaque na análise aos diferentes substantivos do texto. Assim, as *tags* representam como a escrita relaciona tais nomes em estilos próprios e descritivos da linguagem. 215 170 220

4. Datasets

175 Para avaliar o método de interpretação, utilizamos 4 datasets, um em português e três em inglês. O uso destes datasets visa identificar se as características linguísticas da escrita, para além dos termos em si, estão alinhadas com a detecção de fake news. Os seguintes datasets foram utilizados:

4.1. Fake.BR

180 O corpus *Fake.BR* [10] é um dataset brasileiro em português coletado de notícias classificadas em verdadeiras ou falsas. As notícias falsas foram coletadas dos sites *Diário do Brasil*, *A Folha do Brasil*, *The Jornal Brasil*, e *Top Five TV*. As notícias foram filtradas para seleção apenas das que apresentam informações completamente falsas. Por outro lado, os artigos verdadeiros foram coletados das agências de notícia *GI*, *Folha de São Paulo*, e *Estadão*. Uma medida lexicográfica foi utilizada para associar a cada notícia falsa a notícia verdadeira mais similar. Por fim, o dataset foi formado contendo vários tópicos como política, TV & celebridades, sociedade & cotidiano, ciência & tecnologia, economia e religião. No total o corpus contém 7000 notícias, sendo 3500 classificadas como falsas e 3500 verdadeiras. 235 240

4.2. Fake News AMT

195 O *Fake News AMT* é um dataset anotado através do *Amazon Mechanical Turk - AMT* com notícias manualmente sumarizadas [14]. Elas contêm artigos de diferentes temas entre esportes, negócios, entretenimento, política, tecnologia e educação de fontes como *ABCNews*, *CNN*, *USAToday*, *NewYorkTimes*, *FoxNews*, *Bloomberg* e *CNET*. Para cada anotação, foi solicitada à *AMT* a produção de uma notícia falsa citando as mesmas entidades. Assim, o dataset contém 480 artigos. Foram 40 notícias verdadeiras coletadas e 40 falsas produzidas para cada domínio, totalizando 240 para cada classe.

4.3. Celebrity

205 O *Celebrity* é um dataset em inglês formado de notícias de pessoas em colunas sociais como atores, cantores ou políticos [14]. As origens dos artigos foram sites de tabloides americanos conhecidos como o *Entertainment Weekly*, *People Magazine* e *RadarOnline*. Os artigos foram coletados em pares, sendo um verdadeiro e um falso para cada notícia. Para avaliar a veracidade de cada boato foi utilizada a plataforma de checagem *GossipCop.com*. No total, o dataset conta com 250 pares de notícias, totalizando 500 artigos. 210

4.4. LIAR

O *LIAR* dataset [16] é um conjunto de 12.800 declarações curtas coletadas do site *POLITIFACT.COM*. Para cada sentença o site disponibiliza fontes, esclarecimentos e contextos, justificando-os. Assim, de acordo com uma graduação os autores manualmente anotaram as sentenças em 6 classes: *true*, *mostly-true*, *half-true*, *barely-true*, *fake* e *pants-on-fire*. Tais declarações foram coletadas em diferentes contextos e mídias como redes sociais, propagandas, debates, campanhas eleitorais e entrevistas em canais de rádio e televisão.

Outra forma de analisar o dataset, proposto em algumas abordagens [11] é a observação das sentenças de forma binária. Neste formato as notícias são classificadas como verdadeiras quando *true*, *barely-true*, *mostly-true* e falsas quando *half-true*, *fake* e *pants-on-fire*.

5. Experimentos

Nos experimentos testamos os 4 datasets, com a mesma abordagem em português e em inglês. Nós utilizamos para classificação 4 métodos diferentes, *Gradient Boosting*, *Random Forest*, *Support Vector Machine* e *Weightless Neural Networks*. Para os métodos de *NLP* em todos os datasets foram utilizados os modelos do *framework spaCy*³.

Para ver a evolução do aprendizado, nós avaliamos cada algoritmo com *10-Fold* estratificado com 10% e 90% do total de amostras para treinamento. Os resultados de cada classificador para os datasets *Fake.BR*, *Celebrity* e *Fake News AMT* são apresentados nas Tabelas 1 e 2.

Tabela 1: Acurácia média dos classificadores com 90% de treino e 10% teste.

Dataset	Classificador			
	RDF	WSD	GDB	SVM
<i>Fake.BR</i>	0,9578	0,8174	0,9714	0,9576
<i>Celebrity</i>	0,7143	0,5378	0,7990	0,6051
<i>Fake News AMT</i>	0,7813	0,5104	0,9042	0,7521

Tabela 2: Acurácia média dos classificadores com 10% de treino e 90% teste.

Dataset	Classificador			
	RDF	WSD	GDB	SVM
<i>Fake.BR</i>	0,9496	0,8031	0,9589	0,9474
<i>Celebrity</i>	0,5944	0,5849	0,6732	0,5215
<i>Fake News AMT</i>	0,6537	0,5843	0,7764	0,6183

Nas Tabelas 1 e 2 os resultados evidenciam, para além da linguagem, diferentes estilos de escrita e a variação do desempenho dos classificadores. Para o dataset *Fake.BR* os resultados preliminares reportam para o uso de *POSTags* resultados de 75% de acurácia. Isso indica que o uso das estruturas sintáticas não haviam sido efetivas para classificação do *Fake.BR*, enquanto foi positivo para datasets como o *Fake News*

³*spaCy* - <http://spacy.io>

AMT e o *Celebrity*[10]. Porém, o uso de janelas de análise da sintaxe e a identificação de nomes com *POSTag+NER* atingiu 97,14% de acurácia média com o *GDB*. Tais valores superam o alcançando com todas as features extraídas, atingindo acurácia de 89%.

Como podemos observar, o *GDB* foi o classificador com melhor desempenho, tanto com dataset reduzido quanto com o dataset completo. Isso significou acurácia média de 90,42% e 77,64% para *Fake News AMT*. Tal ganho se reflete também no dataset *Celebrity*. O *GDB* apresentou 78,27% e 67,01%. Estes valores são superiores aos 74% de acurácia apresentados pelos autores para *Fake News AMT* e 76% no *Celebrity* [14].

Em geral, o ganho de qualidade na classificação em relação ao acréscimo de informação foi notável, principalmente com os classificadores *GDB* e *RDF*. Para o *GDB*, classificador de maior qualidade para os três datasets, os ganhos foram de 1%, 12% e 13% para os datasets *Fake.BR*, *Celebrity* e *Fake News AMT* respectivamente. Esse mesmo ganho para o *RDF* foi de 1%, 12% e 13% enquanto 1%, 8% e 13% para o *SVM*. Outro aspecto porém ocorre com a *ClusWisard*. Esse classificador apresenta bons resultados com pouca informação porém sem ganhos de desempenho significativos ou até queda de qualidade com o acréscimo de amostras.

O desempenho é relevante também em diferentes estilos de texto, escrita e nível de formalidade. Diferenciando-se da análise de notícias públicas e de amplo alcance, o dataset *Fake News AMT* apresenta *fake news* de criação pessoal de notícias, falsas. Inclusive, é relevante verificar o potencial de avaliação deste método para identificação linguística própria e suas estruturas específicas. Para isso, o *LIAR* dataset, com registros de pequenas sentenças e falas políticas torna-se também uma referência única para validação do modelo. A Tabela 3 apresenta o desempenho dos classificadores neste diferente cenário e com duas classificações distintas gradual (6 classes) e binária (2 classes).

Tabela 3: Avaliação do LIAR dataset com 2 e 6 classes.

Validação				
Dataset	Classificador			
	RDF	WSD	GDB	SVM
<i>LIAR - 6 Classes</i>	0.2274	0.2173	0.2375	0.2383
<i>LIAR - 2 Classes</i>	0.5888	0.5615	0.5818	0.5919
Teste				
Dataset	Classificador			
	RDF	WSD	GDB	SVM
<i>LIAR - 6 Classes</i>	0.2486	0.2044	0.2344	0.2534
<i>LIAR - 2 Classes</i>	0.5959	0.5509	0.5943	0.6054

Uma exceção no uso da estratificação das amostras foi feita para o *LIAR* dataset, onde utilizamos a separação indicada pelo dataset. Nos resultados descritos na Tabela 3, destaca-se a complexidade de graduar níveis de erro. O método de análise puramente linguística e estrutural caracteriza a escrita de um artigo enganoso para além do uso da fonte de origem ou motivações como apresentadas nesse dataset. Os resultados porém, são similares aos apresentados pela literatura [11]. Destaca-se ainda

que, nesse cenário o *SVM* teve desempenho superior, alcançando acurácia de 60,54% e 25,34% para 2 e 6 classes.

6. Conclusão

As *fake news* são problemas na disseminação de informações falsas em massa nos meios de comunicação. São diretamente associadas às redes sociais, aplicativos de mensagem e sites independentes. A credibilidade da notícia porém, em uma visão geral, independe da mídia de divulgação mas sim é relacionada ao conteúdo e escrita características, geralmente com tom ofensivo.

Portanto, neste artigo investigamos se a observação das características estruturais da linguagem tornam o classificador passível de estabelecer se uma notícia é falsa ou não. Para verificação desse estilo, os resultados destacam o uso de estruturas de janela sintática com *POSTag* associado com o *NER*. A classificação foi expressiva e apresenta resultados equivalentes a métodos com amplo uso de informações e metadados. Nessa perspectiva, o *POSTag+NER* foi equivalente à extração da frequência dos termos, sem que os termos isolados em si caracterizem uma notícia falsa ou verdadeira.

Deste modo, podemos considerar que os resultados com a verificação sintática caracteriza as *fake news* com estilos de escrita próprios. Isso as diferencia da descrição factual das agências jornalísticas e as torna passíveis desse tipo de análise. Como trabalhos futuros, o comportamento da estrutura textual entre os datasets e a caracterização linguística das mesmas devem ser analisadas com atenção. Isso auxiliará a ampliar a abrangência dos classificadores na detecção das principais estruturas e verificar como esse comportamento reflete em linguagem natural.

Referências

- [1] Y. Tsifti, H. G. Boomgaarden, J. Strömbäck, R. Vliegenthart, A. Damstra, E. Lindgren, Causes and consequences of mainstream media dissemination of fake news: literature review and synthesis, *Annals of the International Communication Association* 44 (2) (2020) 157–173.
- [2] C. Shao, G. L. Ciampaglia, O. Varol, K.-C. Yang, A. Flammini, F. Menczer, The spread of low-credibility content by social bots, *Nature communications* 9 (1) (2018) 1–9.
- [3] C. T. Robertson, R. R. Mourão, E. Thorson, Who uses fact-checking sites? the impact of demographics, political antecedents, and media use on fact-checking site awareness, attitudes, and behavior, *The International Journal of Press/Politics* 25 (2) (2020) 217–237.
- [4] J. a. Moreno, G. Bressan, Factck.br: A new dataset to study fake news, in: *Proceedings of the 25th Brazilian Symposium on Multimedia and the Web, WebMedia '19*, Association for Computing Machinery, New York, NY, USA, 2019, p. 525–527. doi:10.1145/3323503.3361698. URL <https://doi.org/10.1145/3323503.3361698>
- [5] C. Sindermann, A. Cooper, C. Montag, A short review on susceptibility to falling for fake political news, *Current Opinion in Psychology* 36 (2020) 44–48.
- [6] L. Bozarth, C. Budak, Toward a better performance evaluation framework for fake news classification, *Proceedings of the International AAAI Conference on Web and Social Media* 14 (1) (2020) 60–71.
- [7] K. Shu, A. Sliva, S. Wang, J. Tang, H. Liu, Fake news detection on social media: A data mining perspective, *SIGKDD Explor. Newsl.* 19 (1) (2017) 22–36.
- [8] S. C. R. Gangireddy, D. P. C. Long, T. Chakraborty, Unsupervised fake news detection: A graph-based approach, in: *Proceedings of the 31st ACM Conference on Hypertext and Social Media, HT '20*, Association

for Computing Machinery, New York, NY, USA, 2020, p. 75–83. doi : 10.1145/3372923.3404783.

URL <https://doi.org/10.1145/3372923.3404783>

- [9] S. Yang, K. Shu, S. Wang, R. Gu, F. Wu, H. Liu, Unsupervised fake news detection on social media: A generative approach, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, 2019, pp. 5644–5651.
- [10] R. A. Monteiro, R. L. S. Santos, T. A. S. Pardo, T. A. de Almeida, E. E. S. Ruiz, O. A. Vale, Contributions to the study of fake news in portuguese: New corpus and automatic detection results, in: A. Villavicencio, V. Moreira, A. Abad, H. Caseli, P. Gamallo, C. Ramisch, H. Gonalo Oliveira, G. H. Paetzold (Eds.), Computational Processing of the Portuguese Language, Springer International Publishing, Cham, 2018, pp. 324–334.
- [11] T. Alhindi, S. Petridis, S. Muresan, Where is your evidence: Improving fact-checking by justification modeling, in: Proceedings of the First Workshop on Fact Extraction and VERification (FEVER), Association for Computational Linguistics, Brussels, Belgium, 2018, pp. 85–90.
- [12] H. Ahmed, I. Traore, S. Saad, Detection of online fake news using n-gram analysis and machine learning techniques, in: I. Traore, I. Woungang, A. Awad (Eds.), Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, Springer International Publishing, Cham, 2017, pp. 127–138.
- [13] V. Rubin, N. Conroy, Y. Chen, S. Cornwell, Fake news or truth? using satirical cues to detect potentially misleading news, in: Proceedings of the Second Workshop on Computational Approaches to Deception Detection, Association for Computational Linguistics, San Diego, California, 2016, pp. 7–17.
- [14] V. P  rez-Rosas, B. Kleinberg, A. Lefevre, R. Mihalcea, Automatic detection of fake news, in: Proceedings of the 27th International Conference on Computational Linguistics, Association for Computational Linguistics, Santa Fe, New Mexico, USA, 2018, pp. 3391–3401.
- [15] H. Q. Abonizio, J. I. de Morais, G. M. Tavares, S. Barbon Junior, Language-independent fake news detection: English, portuguese, and spanish mutual features, Future Internet 12 (5) (2020) 87.
- [16] W. Y. Wang, “liar, liar pants on fire”: A new benchmark dataset for fake news detection, in: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), Association for Computational Linguistics, Vancouver, Canada, 2017, pp. 422–426.