Factify 2: A Multimodal Fake News and Satire News Dataset

S Suryavardan*[1], Shreyash Mishra*[1], Parth Patwa[2], Megha Chakraborty[3], Anku Rani[3], Aishwarya Reganti[4], Aman Chadha†[5,6], Amitava Das[3], Amit Sheth[3], Manoj Chinnakotla[7], Asif Ekbal[8] and Srijan Kumar[9]

[1]*IIIT Sri City, India*
[2]*UCLA, USA*
[3]*University of South Carolina, USA*
[4]*Carnegie Mellon University, USA*
[5]*Stanford, USA*
[6]*Amazon AI, USA*
[7]*Microsoft, USA*
[8]*IIT Patna, India*
[9]*Georgia Tech, USA*

Abstract

The internet gives the world an open platform to express their views and share their stories. While this is very valuable, it makes fake news one of our society's most pressing problems. Manual fact checking process is time consuming, which makes it challenging to disprove misleading assertions before they cause significant harm. This is he driving interest in automatic fact or claim verification. Some of the existing datasets aim to support development of automating fact-checking techniques [1, 2], however, most of them are text based. Multi-modal fact verification has received relatively scant attention. In this paper, we provide a multi-modal fact-checking dataset called FACTIFY 2, improving Factify 1 by using new data sources and adding satire articles. Factify 2 has 50,000 new data instances. Similar to FACTIFY 1.0, we have three broad categories - support, no-evidence, and refute, with sub-categories based on the entailment of visual and textual data. We also provide a BERT and Vison Transformer based baseline, which acheives 65% F1 score in the test set. The baseline codes and the dataset will be made available at https://github.com/surya1701/Factify-2.0.

Keywords

Fake News, Fact Verification, Multimodality, Dataset, Machine Learning, Entailment

1. Introduction

With social media platforms taking center stage as news mediums, shifting facts from fake news has become a cause for concern. Fake news articles typically manifest as fabricated stories

*Equal contribution.
†Work does not relate to position at Amazon.

De-Factify 2: 2nd Workshop on Multimodal Fact Checking and Hate Speech Detection, co-located with AAAI 2023. 2023 Washington, DC, USA

suryavardan.s19@iiits.in (S. Suryavardan*); shreyash.m19@iiits.in (S. Mishra*); amitava@mailbox.sc.edu (A. Das)

with no verifiable facts, sources, or quotes. Sometimes these stories may be propaganda that is intentionally designed to mislead the reader or may be designed as "clickbait" written for economic incentives. The technological ease of copying, pasting, clicking, and sharing content online has helped misinformation and disinformation to proliferate. This has causes several challenges in events like Covid-19 [3, 4, 5], elections [6] etc. In some cases, stories are designed to provoke an emotional response and placed on certain sites to entice readers into sharing them widely. In other cases, "fake news" articles may be generated and disseminated by "bots" - computer algorithms that are designed to act like people sharing information, but can do so quickly and automatically [7]. Although there are a few large-scale efforts to identify fake news, like FEVER [1] and LIAR[2], these datasets do not account for the evolution of fake news in the real world. Another hindrance to fake news detection on social media platforms is the fact that online information is very diverse, covering a large number of subjects, which contributes complexity of this task. Often times, the truth and intent of any statement are challenging to be verified by computers alone, so efforts must depend on collaboration between humans and technology, à la human-in-the-loop setting [8]. Additionally, the visual cues that support textual claims would help the system to detect fake content with greater confidence. These concerns were addressed in the previous iteration - FACTIFY 1, which released a multimodal fact-checking dataset for multimodal fact verification. The dataset contains images, textual claims, and reference textual documents/images. It proposed a multimodal entailment task to tag these claims against the verified document/image using 3 classes, i.e., support, no-evidence, and refute; each of these categories is explained in the next section. The first two categories are further sub-divided into text and multimodal components. Thus, in total, all the data samples are labeled with one out of five choices. The data was obtained from twitter handles of popular news channels from two large nations – the US and India. Factify 2 is the latest iteration of factify, where we release new data of 50k instances including satirical articles, which utilize a different manner of presentation of fake news.

The paper is organized as follows: Related work is described in section 2. The proposed task is described in section 3. Data collection and data distribution are explained in section 4 while section 5 demonstrates the baseline model. Section 6 shows the results of our baseline models. Finally, we summarise our task along with the further scope and open-ended pointers in section 7.

2. Related Work

Text based dataset: In recent years, a number of textual datasets for fact-checking and fact-verification have been released. The LIAR [2] dataset contains 13k statements from politiFact [9] annotated into 6 fine-grained labels. FEVER provides manually updated 185k instances of Wikipedia claims and associated supporting documents, categorised as `Support`, `Refute`, or `NotEnoughInfo`. Patwa et al. [10] released a dataset of 10k tweets/articles on Covid-19 annotated as true or false. A dataset for evidence extraction, document retrieval, stance detection, and claim validation is proposed in [11]. [12] create a dataset to differentiate fake news from satire. The PUBHEALTH [13] data has 12k public health claims along with explanations by journalists to support the fact-check labels. Other datasets include [14, 15, 16, 17]. Common methods to

detect text based fake news involve the use of CNN [18], RNN [19, 20], BERT [21, 22, 23], etc.

Multimodal datasets: Text-only databases are inadequate in the social media era. It is crucial to go beyond and consider additional modalities like image and video to detect fake news. The fakeddit [24] dataset contains one million text+image instances taken from reddit and labeled into 6 fine-grained classes for fake news detection. FakeNewsNet provides spatio-temporal and visual data along with news and social context for analysing and detecting fake news. It contains twitter user data such as location, replies, retweets, timestamps, etc. for about 20k multimodal articles from PolitiFact and GossipCop. A multimodal fact-checking dataset called MOCHEG [25] consists of 21,184 assertions, each of which is given a veracity label (`support`, `refute`, and `not enough information`) and an explanation statement. A video dataset consisting 180 verfied and 200 debunked videos is provided by [26]. Some other datasets are [27, 28, 29].

Modelling approaches to this task are varied and unique in their use of classifiers, adversarial training, attention, etc. SpotFake [30] derives textual and visual representations from BERT and VGG, respectively, before concatenating them for classification. EANN [31] trains a fake news classifier adversarially by adding a event discriminator that ensures that the input data is event-invariant such that newly emerging events can also be verified. CARM-N [32] proposes a multichannel nonvolutional neural network that can mitigate the influence of noise information which may be generated by crossmodal attention fusion by extracting textual feature representation from original data and fused textual information simultaneously. Other methods include use of BERT-based CapsNet [33], Cross-modal similarity [34] and Variational Autoencoders [35] among others [36, 37, 38, 39].

Factify 1: FACTIFY [40], is one of the largest multimodal fact-verification public datasets, which includes 50k data points and covers news from India and the US. Images, texts, and reference texts are all part of FACTIFY. They are categorised into three primary groups: `Support`, `Insufficient`, and `Refute`, with additional groups dependent on the inclusion of visual and textual data. FACTIFY 2 follows a similar pattern and releases additional 50k instances which incorporate data from satirical articles and new data sources.

For factify 1, researchers used methods like BERT [41], RoBERTa [42], and BigBird [43] for textual features and ResNet [43], DeiT [44], EfficientNet [45], and VGG [42] for visual features. Please refer to [46] for details of all the methods.

3. The Factify Task

Fact verification is a difficult task to completely automate, especially in the case of multimodal data, given the inherent challenges in doing a holistic evaluation with both the vision and text modalities to ascertain the veracity of the claim. To this end, we model fake news verification as a multimodal entailment task such that the veracity of both the text and image is verified.

The formulation of the task is similar to the previously presented Factify 1. Each sample contains a claim that has to be verified or fact checked. Each claim is accompanied by a supporting document that is to be used to determine the veracity through a comparison or entailment based approach. The claim and document are multi-modal i.e. they have textual and visual data enabling multi-modal entailment for fact verification. Each sample has pairs of text, image and OCR.

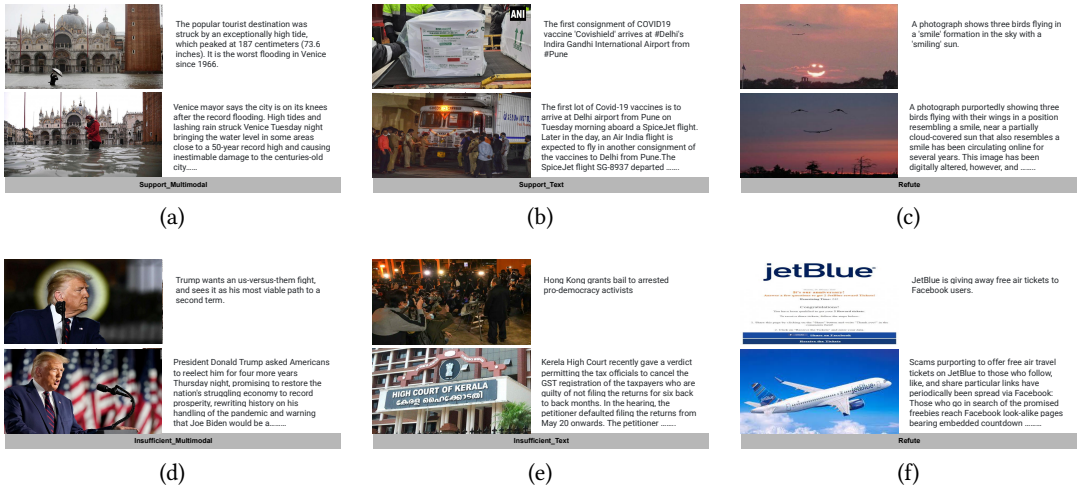


Figure 1: These are examples for all the 5 categories. The document text supports the claim text in images (a) and (b), it is insufficient in images (d) and (e), while it refutes the claim in images (c) and f). The claim and document images are entailed in images (a) and (d) and not entailed in images (b) and (e).

We define the following five categories to describe the entailment of the claim and document: `Support_Text`, `Support_Multimodal`, `Insufficient_Text`, `Insufficient_Multimodal`, and `Refute`. The specific description of these categories is as follows:

- `Support_Text`: the textual data for the claim and document are entailed but their images are not entailed.
- `Support_Multimodal`: the textual data is entailed and the images are also similar for the claim and document.
- `Insufficient_Text`: the textual data is not entailed but the claim and document may have several common words, and the images are not entailed.
- `Insufficient_Multimodal`: the claim and document text are not entailed but they may have common words and the images are also entailed in this case.
- `Refute`: The document text and image both contradict or refute the claim text and image, thus, indicating that the given claim is false.

Some examples from the dataset are given in Figure 1.

4. Data

In this section, we describe the data collection and data analysis.

4.1. Data Collection

The collection process includes two separate pipelines: (i) to collect real news articles for support and no-evidence classes, and (ii) to collect fake news articles for no-evidence and refute

classes. The end goal was to curate a dataset with text and image for both claims and their corresponding supporting documents.

The first part of the collection was similar to FACTIFY 1. We collected tweets date-wise from renowned twitter news handles, namely Hindustan Times, ANI and ABC, CNN for India and USA, respectively. The nature and format of these handles as well as their tweets aided our objective of collecting real news claims and articles. To improve the diversity and functionality of the dataset, we compared tweets across the news handles to identify tweets that were reporting the same or similar news. For this, we followed steps similar to FACTIFY 1, where we compare tweet texts using Sentence BERT [47] and using a threshold we categorise whether it is the same news or not. Specifically, we use the pre-trained `paraphrase-MiniLM-L6-v2` [48] variant of Sentence BERT (SBERT) [49] instead of alternatives such as BERT or RoBERTa, owing to its rich sentence embeddings yielding superior performance [47], while being much more time-efficient. If the news is not the same, we compare common words using the NLTK library [50] to categorise the tweets as similar or dissimilar. This helps define the support and no-evidence category respectively as described in Section 3. The similarity between images in the compared tweet pairs are also used to further categorise the data based on visual entailment. Thresholds were set for image similarity to categorise them as entailed or not, based on two metrics: cosine similarity between ResNet50 embeddings and Histogram similarity. With this collected data, we treated the tweet from one handle as the claim and the news article associated with the tweet from the other handle as the supporting document.

The second part is the collection from several different websites. A part of the data for refute category was collected from fact checking websites, similar to FACTIFY 1. We scraped data from Snopes [51], Factly [52] and Boom [53]. These websites provided a well-defined claim and a document disproving the given claim. We added an additional data source in this iteration of the task, we collected satirical articles that were fake in nature but were written in a way that seems real to the reader. While the websites we scraped from i.e. Fauxy [54] and EmpireNews [55], specify that their articles are not true, we added them to the support category. This is because, as aforementioned, the articles support their claim despite the claim being fake in nature. To make the claim multi-modal, we scraped images by searching for the headline of the article. We also manually annotated some articles we collected from the search results of these headlines to add data to the no-evidence and refute category in cases where the articles were about these satirical claims.

4.2. Data Statistics And Analysis

The second iteration of FACTIFY has the same categories as FACTIFY 1, with 50,000 data samples. The samples are equally divided among all five categories with a split of 70:15:15 into train, validation, and test sets respectively.

Key words can be vital when identifying or predicting the veracity of a given claim. By analyzing the claim and their documents, we find the most frequently occurring words in Figure 2. Most of the words relate to politics, indicating the bias in the news articles.

The political inclination of the dataset is re-iterated by the word cloud for the support and no-evidence category in Figure 3. However, in the same image, the refute category has a more general distribution of words, with several words related to social media present in the refute

	Train	Validation	Test	Total
Support_Multimodal	7000	1500	1500	10000
Support_Text	7000	1500	1500	10000
Insufficient_Multimodal	7000	1500	1500	10000
Insufficient_Text	7000	1500	1500	10000
Refute	7000	1500	1500	10000
Total	35000	7500	7500	**50000**

Table 1
Dataset distribution statistics for the FACTIFY 2 dataset. Note that the data is balanced across categories.

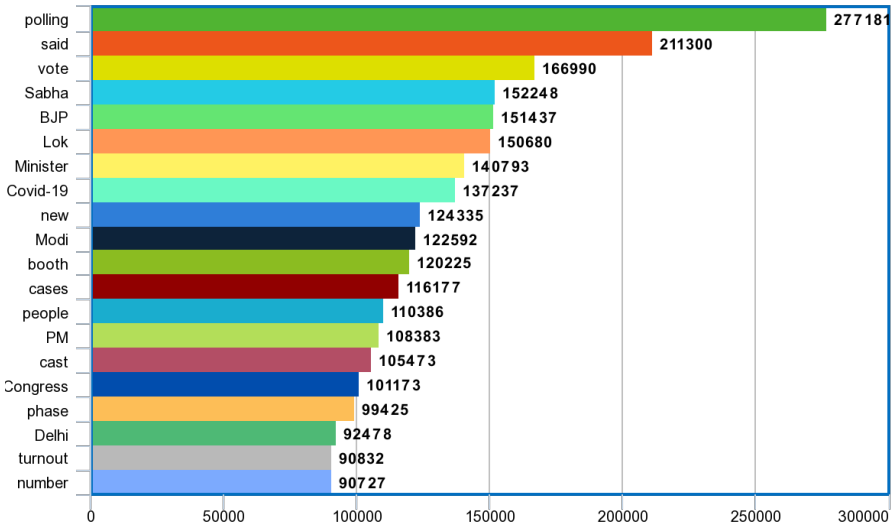


Figure 2: The top 20 most frequent words extracted from the claim documents and their frequencies. Many of the words are related to politics.

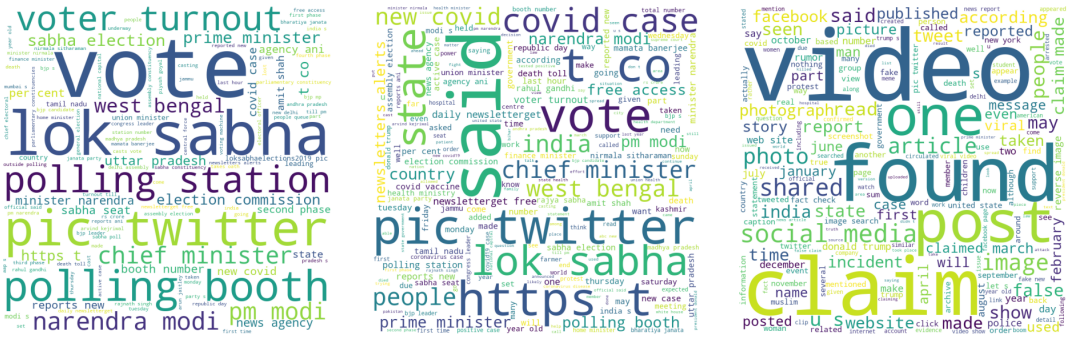
N-gram	Examples
1-gram	(polling), (vote)
2-gram	(lok,sabha), (polling,booth)
3-gram	(Prime,Minister,Narendra), (access,to,newsletters)
4-gram	(Prime,Minister,Narendra,Modi), (phase,of,Lok,Sabha)

Table 2
N-gram examples for the claims of all categories.

category word cloud. We further present unique n-gram examples for the FACTIFY 2 dataset in Table 2 to show the lexical diversity of the dataset.

5. Baseline model

Several media are regularly used in online information exchange. Pictures have the power to misrepresent a claim and propagate erroneous information. We must consider both the



(a) Support (b) Insufficient (c) Refute

Figure 3: Word clouds indicating top words used in each class. Words related to politics and Covid dominate in the support and insufficient categories.

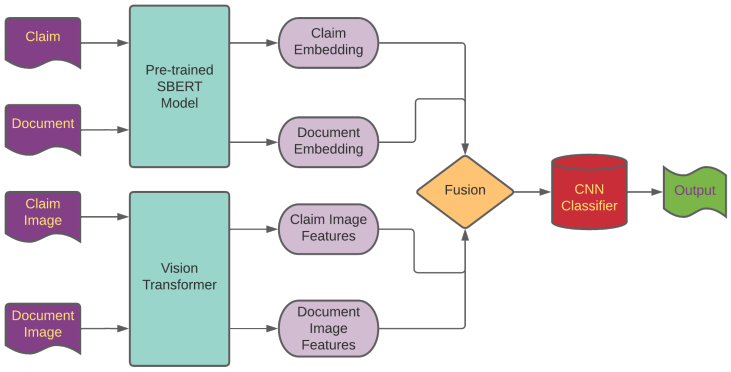


Figure 4: Baseline model architecture. Text, image features extracted from the document and the claim are concatenated and used for final prediction.

image and the text in order to appropriately classify the claims. Features must be obtained from claim and document image-text pairings because it is an entailment-based technique. The visual features are obtained from the pre-trained Vision transformer model (ViT) [56]. Thanks to the positional embedding of picture patches carried out by ViT, the ViT model can surpass conventional CNNs in terms of computation and accuracy. Using a pretrained Sentence BERT model (specifically, the `stsb-mpnet-base-v2` variant), the model generates sentence embeddings of claim and document attributes. The Sentence-BERT embedding is concatenated with the pooled output from the ViT model. After passing through an MLP, the combined features are then categorised. The multi-modal characteristics are employed for all three of the sub-tasks after modifications to the MLP. The model architecure is displayed in Figure 4. The codes will be made available at https://github.com/surya1701/Factify-2.0.

6. Results

Baseline results in Table 3 show Macro F1 scores for some muti-modal modelling approaches mentioned below. Using ViT for extracting visual features and Sentence-BERT for the textual features, the baseline model scores 0.6499. We also compare the baseline model (ViT + SBERT-MPNet) with other methods, such as ViT + SBERT-RoBERTa, in which the SBERT-RoBERTa model is used in place of SBERT-RoBERTa for generating text embeddings. For the Resnet50 + SBERT-RoBERTa and Resnet50 + SBERT-MPNet, a simple ResNet50 model is used to extract visual features. The improvement on using the Vision transformer over the ResNet model signifies the importance of images for the task.

Method	Macro F1
Resnet50 + SBERT-RoBERTa	0.4504
Resnet50 + SBERT-MPNet	0.4727
ViT + SBERT-RoBERTa	0.6226
ViT + SBERT-MPNet	**0.6499**

Table 3
Baseline scores on the test set. ViT based models significantly outperform resnet based models.

7. Conclusion and Future Work

By publishing a sizable real-world dataset containing inputs from two modalities, namely text and image, we make a significant step towards creating machine learning approaches for the multimodal fact verification in this study. To underline the difficulties of the issue and the scope for improvement, we conduct data analysis and release multimodal baselines. However, there are a lot of additional research possibilities that can be explored since our work merely touches the surface. One potential research direction could be to enrich the dataset with reasoning that why is a particular news fake. Another possibility is to use synthetic data that matches the general data distribution, thus adding complexity to the refute category.

References

[1] J. Thorne, A. Vlachos, C. Christodoulopoulos, A. Mittal, Fever: a large-scale dataset for fact extraction and verification, arXiv preprint arXiv:1803.05355 (2018).

[2] W. Y. Wang, " liar, liar pants on fire": A new benchmark dataset for fake news detection, arXiv preprint arXiv:1705.00648 (2017).

[3] N. Karimi, J. Gambrell, hundreds die of poisoning in iran as fake news suggests methanol cure for virus, 2020. URL: https://www.timesofisrael.com/hundreds-die-of-poisoning-in-iran-as-fake-news-suggests-methanol-cure-for-virus/.

[4] J. Bae, D. Gandhi, J. Kothari, S. Shankar, J. Bae, P. Patwa, R. Sukumaran, A. Chharia, S. Adhikesaven, S. Rathod, I. Nandutu, S. TV, V. Yu, K. Misra, S. Murali, A. Saxena, K. Jakimowicz, V. Sharma, R. Iyer, A. Mehra, A. Radunsky, P. Katiyar, A. James, J. Dalal, S. Anand, S. Advani, J. Dhaliwal, R. Raskar, Challenges of equitable vaccine distribution in the covid-19 pandemic, 2022. arXiv:2012.12263.

[5] M. Morales, R. Barbar, D. Gandhi, S. Landage, J. Bae, A. Vats, J. Kothari, S. Shankar, R. Sukumaran, H. Mathur, K. Misra, A. Saxena, P. Patwa, S. T. V., M. Arseni, S. Advani, K. Jakimowicz, S. Anand, P. Katiyar, A. Mehra, R. Iyer, S. Murali, A. Mahindra, M. Dmitrienko, S. Srivastava, A. Gangavarapu, S. Penrod, V. Sharma, A. Singh, R. Raskar, Covid-19 tests gone rogue: Privacy, efficacy, mismanagement and misunderstandings, 2021. arXiv:2101.01693.

[6] S. Muhammed T, S. K. Mathew, The disaster of misinformation: a review of research in social media, International Journal of Data Science and Analytics 13 (2022) 271–285.

[7] S. Desai, H. Mooney, J. A. Oehrli, Fake news, lies and propaganda: How to sort fact from fiction, Subjects: News & Current Events. The University of Michigan Library (2021).

[8] X. Zhang, A. A. Ghorbani, An overview of online fake news: Characterization, detection, and discussion, Information Processing & Management 57 (2020) 102025.

[9] Politifact, https://www.politifact.com, 2007. Accessed: 2022.

[10] P. Patwa, S. Sharma, S. Pykl, V. Guptha, G. Kumari, M. S. Akhtar, A. Ekbal, A. Das, T. Chakraborty, Fighting an infodemic: Covid-19 fake news dataset, in: Combating Online Hostile Posts in Regional Languages during Emergency Situation (CONSTRAINT) 2021, Springer, 2021, p. 21–29. URL: http://dx.doi.org/10.1007/978-3-030-73696-5_3. doi:10.1007/978-3-030-73696-5_3.

[11] A. Hanselowski, C. Stab, C. Schulz, Z. Li, I. Gurevych, A richly annotated corpus for different tasks in automated fact-checking, in: Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL), Association for Computational Linguistics, Hong Kong, China, 2019, pp. 493–503. URL: https://aclanthology.org/K19-1046. doi:10.18653/v1/K19-1046.

[12] J. Golbeck, M. Mauriello, B. Auxier, K. H. Bhanushali, C. Bonk, M. A. Bouzaghrane, C. Buntain, R. Chanduka, P. Cheakalos, J. B. Everett, W. Falak, C. Gieringer, J. Graney, K. M. Hoffman, L. Huth, Z. Ma, M. Jha, M. Khan, V. Kori, E. Lewis, G. Mirano, W. T. Mohn IV, S. Mussenden, T. M. Nelson, S. Mcwillie, A. Pant, P. Shetye, R. Shrestha, A. Steinheimer, A. Subramanian, G. Visnansky, Fake news vs satire: A dataset and analysis, in: Proceedings of the 10th ACM Conference on Web Science, WebSci '18, Association for Computing Machinery, New York, NY, USA, 2018, p. 17–21. URL: https://doi.org/10.1145/3201064.3201100.

doi:10.1145/3201064.3201100.
[13] N. Kotonya, F. Toni, Explainable automated fact-checking for public health claims, in: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), Association for Computational Linguistics, Online, 2020, pp. 7740–7754. URL: https://aclanthology.org/2020.emnlp-main.623. doi:10.18653/v1/2020.emnlp-main.623.
[14] R. Mihalcea, C. Strapparava, The lie detector: Explorations in the automatic recognition of deceptive language, in: Proceedings of the ACL-IJCNLP 2009 Conference Short Papers, ACLShort '09, Association for Computational Linguistics, USA, 2009, p. 309–312.
[15] A. Kazemi, K. Garimella, D. Gaffney, S. A. Hale, Claim matching beyond english to scale global fact-checking, 2021. arXiv:2106.00853.
[16] T. Mitra, E. Gilbert, Credbank: A large-scale social media corpus with associated credibility annotations, in: ICWSM, 2015.
[17] I. Augenstein, C. Lioma, D. Wang, L. Chaves Lima, C. Hansen, C. Hansen, J. Grue Simonsen, Multifc: A real-world multi-domain dataset for evidence-based fact checking of claims, in: EMNLP, Association for Computational Linguistics, 2019.
[18] H. Saleh, A. Alharbi, S. H. Alsamhi, Opcnn-fake: Optimized convolutional neural network for fake news detection, IEEE Access 9 (2021) 129471–129489. doi:10.1109/ACCESS.2021.3112806.
[19] O. Ajao, D. Bhowmik, S. Zargari, Fake news identification on twitter with hybrid cnn and rnn models, in: Proceedings of the 9th International Conference on Social Media and Society, SMSociety '18, Association for Computing Machinery, New York, NY, USA, 2018, p. 226–230. URL: https://doi.org/10.1145/3217804.3217917. doi:10.1145/3217804.3217917.
[20] J. A. Nasir, O. S. Khan, I. Varlamis, Fake news detection: A hybrid cnn-rnn based deep learning approach, International Journal of Information Management Data Insights 1 (2021) 100007. URL: https://www.sciencedirect.com/science/article/pii/S2667096820300070. doi:https://doi.org/10.1016/j.jjimei.2020.100007.
[21] R. K. Kaliyar, A. Goswami, P. Narang, Fakebert: Fake news detection in social media with a bert-based deep learning approach, Multimedia tools and applications 80 (2021) 11765–11788.
[22] P. Patwa, M. Bhardwaj, V. Guptha, G. Kumari, S. Sharma, S. PYKL, A. Das, A. Ekbal, S. Akhtar, T. Chakraborty, Overview of constraint 2021 shared tasks: Detecting english covid-19 fake news and hindi hostile posts, in: Proceedings of the First Workshop on Combating Online Hostile Posts in Regional Languages during Emergency Situation (CONSTRAINT), Springer, 2021.
[23] A. Glazkova, M. Glazkov, T. Trifonov, g2tmn at constraint@AAAI2021: Exploiting CT-BERT and ensembling learning for COVID-19 fake news detection, in: Combating Online Hostile Posts in Regional Languages during Emergency Situation, Springer International Publishing, 2021, pp. 116–127. doi:10.1007/978-3-030-73696-5_12.
[24] K. Nakamura, S. Levy, W. Y. Wang, r/fakeddit: A new multimodal benchmark dataset for fine-grained fake news detection, arXiv preprint arXiv:1911.03854 (2019).
[25] B. M. Yao, A. Shah, L. Sun, J.-H. Cho, L. Huang, End-to-end multimodal fact-checking and explanation generation: A challenging dataset and models, arXiv preprint arXiv:2205.12487 (2022).

[26] O. Papadopoulou, M. Zampoglou, S. Papadopoulos, Y. Kompatsiaris, A corpus of debunked and verified user-generated videos, Online Information Review 43 (2019) 72–88.
[27] J. C. S. Reis, P. de Freitas Melo, K. Garimella, J. M. Almeida, D. Eckles, F. Benevenuto, A dataset of fact-checked images shared on whatsapp during the brazilian and indian elections, 2020. arXiv:2005.02443.
[28] S. Jindal, R. Sood, R. Singh, M. Vatsa, T. Chakraborty, Newsbag: A multimodal benchmark dataset for fake news detection, in: CEUR Workshop Proc., volume 2560, 2020, pp. 138–145.
[29] D. Zlatkova, P. Nakov, I. Koychev, Fact-checking meets fauxtography: Verifying claims about images, 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), Association for Computational Linguistics, Hong Kong, China, 2019, pp. 2099–2108. URL: https://aclanthology.org/D19-1216. doi:10.18653/v1/D19-1216.
[30] S. Singhal, R. R. Shah, T. Chakraborty, P. Kumaraguru, S. Satoh, Spotfake: A multi-modal framework for fake news detection, in: 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM), 2019, pp. 39–47. doi:10.1109/BigMM.2019.00-44.
[31] Y. Wang, F. Ma, Z. Jin, Y. Yuan, G. Xun, K. Jha, L. Su, J. Gao, Eann: Event adversarial neural networks for multi-modal fake news detection, 2018, pp. 849–857. doi:10.1145/3219819.3219903.
[32] C. Song, N. Ning, Y. Zhang, B. Wu, A multimodal fake news detection model based on cross-modal attention residual and multichannel convolutional neural networks, Information Processing & Management 58 (2020). doi:10.1016/j.ipm.2020.102437.
[33] B. Palani, S. Elango, V. Viswanathan K, Cb-fake: A multimodal deep learning framework for automatic fake news detection using capsule neural network and bert, Multimedia Tools Appl. 81 (2022) 5587–5620. URL: https://doi.org/10.1007/s11042-021-11782-3. doi:10.1007/s11042-021-11782-3.
[34] X. Zhou, J. Wu, R. Zafarani, Safe: Similarity-aware multi-modal fake news detection, in: Advances in Knowledge Discovery and Data Mining: 24th Pacific-Asia Conference, PAKDD 2020, Singapore, May 11–14, 2020, Proceedings, Part II, Springer, 2020, pp. 354–367.
[35] D. Khattar, J. Singh, M. Gupta, V. Varma, Mvae: Multimodal variational autoencoder for fake news detection, 2019, pp. 2915–2921. doi:10.1145/3308558.3313552.
[36] V. Pérez-Rosas, B. Kleinberg, A. Lefevre, R. Mihalcea, Automatic detection of fake news, 2017. arXiv:1708.07104.
[37] J. Ma, W. Gao, K.-F. Wong, Rumor detection on Twitter with tree-structured recursive neural networks, in: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Melbourne, Australia, 2018, pp. 1980–1989. URL: https://aclanthology.org/P18-1184. doi:10.18653/v1/P18-1184.
[38] S. R. Sahoo, B. B. Gupta, Multiple features based approach for automatic fake news detection on social networks using deep learning, Applied Soft Computing 100 (2021) 106983.
[39] Z. Guo, M. Schlichtkrull, A. Vlachos, A survey on automated fact-checking, Transactions of the Association for Computational Linguistics 10 (2022) 178–206.
[40] S. Mishra, S. Suryavardan, A. Bhaskar, P. Chopra, A. Reganti, P. Patwa, A. Das,

T. Chakraborty, A. Sheth, A. Ekbal, et al., Factify: A multi-modal fact verification dataset, in: Proceedings of the First Workshop on Multimodal Fact-Checking and Hate Speech Detection (DE-FACTIFY), 2022.
[41] A. Dhankar, O. Zaiane, F. Bolduc, Uofa-truth at factify 2022: A simple approach to multi-modal fact-checking (2022).
[42] Y. Zhuang, Y. Zhang, Yet at factify 2022: Unimodal and bimodal roberta-based models for fact checking, in: Proceedings of De-Factify: Workshop on Multimodal Fact Checking and Hate Speech Detection, CEUR, 2022.
[43] J. Gao, H.-F. Hoffmann, S. Oikonomou, D. Kiskovski, A. Bandhakavi, Logically at factify 2022: Multimodal fact verification (2022).
[44] W.-Y. Wang, W.-C. Peng, Team yao at factify 2022: Utilizing pre-trained models and co-attention networks for multi-modal fact verification (2022).
[45] N. Hulke, B. R. Siva, A. Raj, A. A. Saifee, Tyche at factify 2022: Fusion networks for multi-modal fact-checking (2021).
[46] P. Patwa, S. Mishra, S. Suryavardan, A. Bhaskar, P. Chopra, A. Reganti, A. Das, T. Chakraborty, A. Sheth, A. Ekbal, et al., Benchmarking multi-modal entailment for fact verification, in: Proceedings of De-Factify: Workshop on Multimodal Fact Checking and Hate Speech Detection, CEUR, 2022.
[47] N. Reimers, I. Gurevych, Sentence-bert: Sentence embeddings using siamese bert-networks, in: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 2019. URL: https://arxiv.org/abs/1908.10084.
[48] paraphrase-minilm-l6-v2, https://huggingface.co/sentence-transformers/paraphrase-MiniLM-L6-v2, 2019. Accessed: 2022.
[49] N. Reimers, I. Gurevych, Sentence-bert: Sentence embeddings using siamese bert-networks, arXiv preprint arXiv:1908.10084 (2019).
[50] S. Bird, E. Klein, E. Loper, Natural language processing with Python: analyzing text with the natural language toolkit, " O'Reilly Media, Inc.", 2009.
[51] Snopes, https://www.snopes.com/, 1994. Accessed: 2022.
[52] Factly, https://factly.in/category/english/, 2016. Accessed: 2022.
[53] Boomlive, https://www.boomlive.in/fact-check, 2014. Accessed: 2022.
[54] Fauxy, https://thefauxy.com/, 2018. Accessed: 2022.
[55] Empirenews, https://empirenews.net/, 2014. Accessed: 2022.
[56] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al., An image is worth 16x16 words: Transformers for image recognition at scale, arXiv preprint arXiv:2010.11929 (2020).