DAT@Z21: A Comprehensive Multimodal Dataset for Rumor Classification in Microblogs

Abderrazek Azri[1][0000−0003−3204−5740], Cécile Favre[1], Nouria Harbi[1], Jérôme Darmont[1][0000−0003−1491−384X], and Camille Noûs[2]

[1] Université de Lyon, Lyon 2, UR ERIC
5 avenue Pierre Mendès France, F69676 Bron Cedex, France
[2] Université de Lyon, Lyon 2, Laboratoire Cogitamus
{a.azri,cecile.favre,nouria.harbi,jerome.darmont}@univ-lyon2.fr
camille.nous@cogitamus.fr

Abstract. Microblogs have become popular media platforms for reporting and propagating news. However, they also enable the proliferation of misleading information that can cause serious damage. Thus, many efforts have been taken to defeat rumors automatically. While several innovative solutions for rumor detection and classification have been developed, the lack of comprehensive and labeled datasets remains a major limitation. Existing datasets are scarce and none of them provide all of the features that have proven to be effective for rumor analysis. To mitigate this problem, we propose a big data-sized dataset called DAT@Z21, which provides news contents with rich features including textual contents, social context, social engagement of users and spatiotemporal information. Furthermore, DAT@Z21 also provides visual contents, i.e., images, which play a crucial role in the news diffusion process. We conduct exploratory analyses to understand our dataset's characteristics and analyze useful patterns. We also experiment various state-of-the-art rumor classification methods to illustrate DAT@Z21's usefulness, especially its visual components. Eventually, DAT@Z21 is available online at https://git.msh-lse.fr/eric/dataz21.

Keywords: Social networks · Rumors · Datasets · Multimodal learning.

1 Introduction

The explosive growth of social media platforms has led to the generation and proliferation of a large amount of data that reaches a wide audience. The openness and unrestricted method of sharing information on microblogging platforms fosters information spread regardless of its credibility. Misinformation, which is also known as rumors, may cause severe damages in the real-world. For example, during the US presidential elections, it was estimated that over 1 million tweets were related to "Pizzagate," a piece of fake news from Reddit, which led to a real shooting[1].

[1] https://en.wikipedia.org/wiki/Pizzagate_conspiracy_theory

To mitigate this problem, automatic rumor detection on social media has recently started to attract considerable attention. Many innovative and significant solutions have been proposed in the literature to defeat rumors (see a comprehensive survey in [2,13]). However, these techniques remain dependent on the availability of comprehensive and community-driven rumor datasets, which have become a major road block.

Most of the publicly available datasets contain only news contents or linguistic features [20,12,14]. Apart from their individual limitations, the common drawback of those datasets is the lack of social context and information other than the text of news articles. In addition to news contents, a few datasets also contain social context, such as user comments and reposts on social media platforms [11,19,16]. Although these engagement-driven datasets are valuable for fake news detection, they have mostly not covered any user profiles, except FakeNewsNet [16].

Detecting rumors on social media is a very challenging task because fake news pieces are intentionally written to mislead consumers, which makes it difficult to spot fake news from news content only. Thus, we need to explore information in addition to news content, such as the social engagements and social behavior of users on social media [16].

The existing datasets cannot address this challenges. Thus, in this study we collect a comprehensive and large-scale repository with multidimensional information, such as textual and visual content, spatiotemporal information, social engagements, and the behaviors of users. The main contributions of this work are summarized as follows:

- We aim to construct and release a comprehensive rumor dataset. DAT@Z21 provides multidimensional information on news articles and social media posts, while our dataset includes rich features such as textual, visual, spatiotemporal, and network information. We provide the methodological details on how it is built.
- We perform various exploratory analyses (data statistics and distributions) on news articles and on the social diffusion (tweets) of the dataset to understand their key properties and characteristics.
- We conduct extensive experiments on the rumor classification task using the DAT@Z21 data and several baselines to validate its usefulness, especially its visual components. Baselines are obtained using either single-modal or multimodal information of tweets, and we utilize either traditional machine learning algorithms or deep learning models.

The rest of this paper is organized as follows. We first survey related works in Section 2. We detail the data collection process, and we present exploratory statistics, analysis and data distribution in Section 3. Experiments using DAT@Z21 dataset to classify rumor credibility are designed and conducted in Section 4. We conclude this paper and shed some light on future work in Section 5.

2 Related Works

In this section, we briefly review the related datasets. These datasets can be grouped into those that were recently designed for fake health news related to the pandemic COVID-19, and datasets for fake news and rumors in general.

2.1 Fake Health News Datasets

With the emergence of the COVID-19 pandemic, political and medical rumors have increased to a level where they have created what is commonly referred to as the global infodemic. In a short amount of time, many COVID-19 datasets have been released. Most of these datasets either have no annotations at all, they employ automated annotations using transfer learning or semi-supervised methods, or are not specifically designed for COVID-19 misinformation.

The earliest and most noteworthy dataset depicting the COVID-19 pandemic at a global scale was contributed by the John Hopkins University [5]. The authors developed an online real-time interactive dashboard[2], which was first made public in January 2020. This dashboard lists the daily number of positive cases, the number of cured patients, and the mortality rates at a country and state/province level.

In terms of datasets collected for COVID-19 infodemic analysis and detection, examples include CoAID [3], which contains automatic annotations for tweets, replies, and claims for fake news; ReCOVery [21], which is a multimodal dataset annotated for tweets sharing reliable versus unreliable news, annotated via distant supervision; FakeCovid [15], which is a multilingual cross-domain fake news detection dataset with manual annotations; and [4], which is a large-scale Twitter dataset that is also focused on fake news. A survey of the different COVID-19 datasets can be found in [9,18].

2.2 Fake News Datasets

A high-quality dataset plays an extremely important role in the task of rumor classification. However, the lack of labeled fake news datasets is a major bottleneck when building an effective detection system for online misleading information. Generally, news data with ground truth are gathered by either expert journalists, fact-checking websites or crowdsourcing. Although, there is no consensual benchmark for fake news detection [17], some publicly available resources are worth mentioning. Most of them only contain news contents, and more particularly textual content information.

The BuzzFeedNews[3] dataset contains URLs from posts produced by nine verified Facebook publishers over a week close to the 2016 U.S. election from September 19 to 23 and September 26 and 27. Each post and linked article is manually fact-checked and annotated for veracity by BuzzFeed journalists.

[2] https://coronavirus.jhu.edu/map.html
[3] https://www.buzzfeed.com/

It contains 1,627 articles–826 mainstream, 356 left-wing, and 545 right-wing articles.

LIAR[4] [20] dataset is collected from fact-checking website PolitiFact[5]. It has 12.8K human labeled short statements collected from PolitiFact and the statements are labeled into six categories ranging from completely false to completely true as pants on fire, false, barely true, half-true, mostly true, and true. It is useful to find misinformation in short statements, but cannot be applied for complete news articles.

The BS Detector[6] is a web crawler with knowledge about fake news websites, which was used to build a dataset by monitoring these websites. Labels are output by the BS detector not by human annotators.

NELA-GT-2018 [12] provides 714k news articles in general topics from 194 news producers. The labels are obtained from eight assessment sites. Among the 194 news sources, 40 of them are not found any labels from the assessments sites.

FA-KES [14] is a fake news dataset that focuses on the Syrian war. It contains 804 articles that labeled as real or fake. However, the labels were annotated based on a database from the Syrian Violations Documentation Center with a cluster algorithm, so the reliability of these labels may be of concerns.

Apart from their own limitations, the common drawback of the data listed here is the lack of social context and information other than the text of news articles. Besides the news content, other researchers have also collected user engagement features of the news, such as user engagement on online social media.

FakeNewsNet[7] [16] is collected from fact-checking websites PolitiFact and GossipCop. News contents and the related ground truth labels are crawled from these two websites. The authors collected social engagement through Twitter's Advanced Search API. This dataset contains 1,056 articles from PolitiFact and 22,864 articles from GossipCop. Each article is labeled as fake or true. Overall, the dataset nearly contains two million tweets .

CREDBANK[8] [11] contains about 1000 news events whose credibility are labeled by 30 annotators who are sourced from Amazon mechanical Turk. This dataset contains 60 million tweets that were posted between 2015 and 2016.

Although these datasets are valuable for fake news detection, they mostly have not cover any user profiles except FakeNewsNet. Moreover, only one dataset includes visual data.

The existing datasets often only focus on the textual content of news articles, and little attention has been paid to visual content of both news articles and user engagement of news on social media. However, fake images on media posts can easily go viral on social platforms, and cause serious social disruptions.

Unlike this category of datasets, we propose in this paper a new dataset: DAT@Z21. It not only includes both true and fake news articles but also their

[4] https://www.cs.ucsb.edu/william/data/liardataset.zip
[5] https://www.politifact.com/factchecks/
[6] https://www.kaggle.com/mrisdal/fake-news
[7] https://github.com/KaiDMML/FakeNewsNet
[8] http://compsocial.github.io/CREDBANK-data/

diffusion on social media. It contains all possible features of interests, including multi-modal information (textual, visual and spatiotemporal information), and user engagement data (user profile, tweets, retweets, replies, likes) as shown in Table 1 that compares our new dataset to existing ones. In addition, DAT@Z21 can be updated automatically to fetch the latest news article and social media content.

Table 1: Comparison with Existing Fake News Datasets

Features	**News Content**		**User Engagement**					**Spatiotemporal data**	
Datasets	textual	Visual	User profile	Tweet	Retweet	Reply	Attached image	Spatial	Temporal
BuzzFeedNews	✓	—	—	—	—	—	—	—	—
LIAR	✓	—	—	—	—	—	—	—	—
BS Detector	✓	—	—	—	—	—	—	—	—
NELA-GT-2018	✓	—	—	—	—	—	—	—	—
FA-KES	✓	—	—	—	—	—	—	—	—
FakeNewsNet	✓	✓	✓	✓	✓	—	—	✓	✓
CREDBANK	✓	—	✓	✓	—	—	—	✓	✓
DAT@Z21	✓	✓	✓	✓	✓	✓	✓	✓	✓

3 Data Collection

In this section, we present the process that we followed to build DAT@Z21. We show how we collected data with a valuable ground truth. We will also show how we collected news articles and messages that were posted on social media. The overall data collection process can mainly be divided into three stages, as shown in Figure 1: (1) collecting news articles with ground truth labels; (2) preparing tweets collection; and (3) collecting content generated on social media. The following subsections will describe these stages in more detail.

3.1 News Articles & Ground Truth Collection

We start by collecting as of August 30, 2021 all verified rumor news articles, with reliable ground truth labels from the fact-checking website *PolitiFact*[9]. PolitiFact is a well-known fact-checking and rumor debunking website that contains many types of reports. For each checked rumor, professional analysts provide a conclusion of the rumor followed by a full description, source, origin, supporting/opposing arguments of the rumor story, as well as the truth label. Each statement is divided into six categories of veracity: true, mostly true, half-true, mostly false and "pants on fire" for outrageously false statements.

We use the source URLs of web pages that publish the news articles, provided by PolitiFact professionals, to crawl the related news contents. After identifying the web pages, we access and retrieve them using Politifact API. We then pull

[9] www.PolitiFact.com/factchecks/

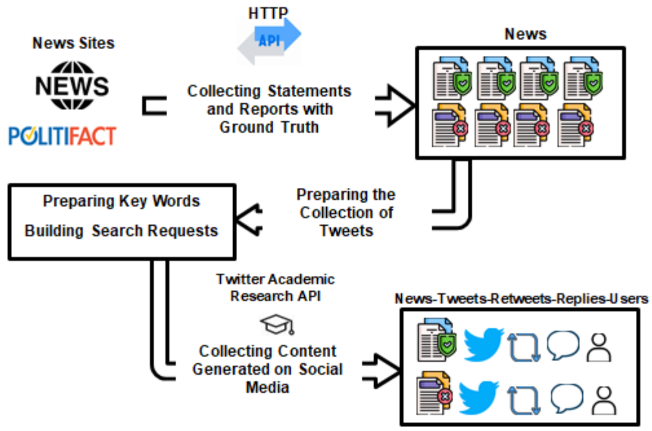


Fig. 1: Data Collection pipeline for DAT@Z21

data out from the content of these pages using *Beautiful Soup Python library*[10]. To guarantee a high quality ground truth, and reduce the number of false positive and false negative in the news labels of our dataset, we select only news articles that are explicitly tagged as true, false, or grossly false. We are particularly interested in the information related to each news article, which are:

1. Article ID: We assign a unique identifier to each news article;
2. Article Title: The title of the news article, which might well summarize or give clues about the content of the article;
3. Article Author: The author(s) of the news article;
4. Publication Date: The date of publication of the article;
5. Article URL: The URL to the news article;
6. Article Body Text: The textual content of the news article;
7. Article Image: The URL to the image(s) attached to the article;
8. Article Label: The original ground truth of the article credibility, either as True or Fake.

Figures 2 and 3 show some of the textual features of news content generated from both the title and the body text. From Figure 2, we can note that the number of words in articles follows a scaling long-tail distribution, with a mean value of $\simeq 2,513$ and an average of $\simeq 2,418$. Furthermore, Figure 3 shows the word cloud of all articles. Because the news articles are statements, reports and claims, which could be potentially a fake or real rumors, they naturally frequently share some vocabularies, such as "said" ($\#= 11,131$), "state" ($\#= 10,112$), "stated" ($\#= 9,373$), and "according" ($\#= 9,208$).

We can also capture the source of the rumor from the titles of the news articles. In PolitiFact, this source is usually a statement from a politician but it

[10] https://pypi.org/project/beautifulsoup4/

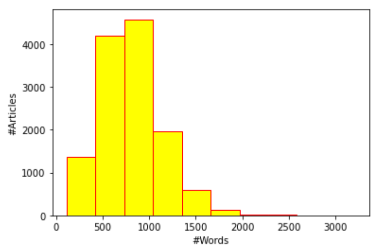


Fig. 2: Word Count Distribution

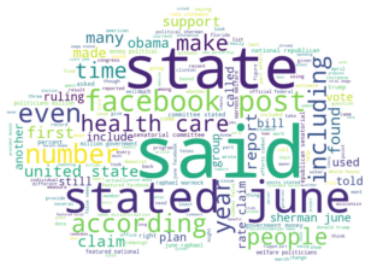
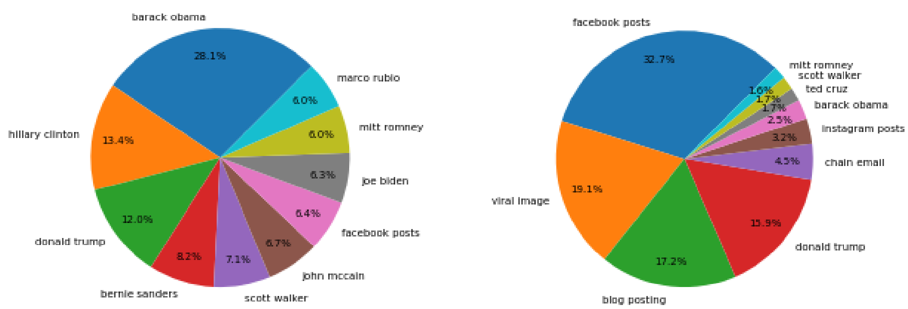


Fig. 3: Word Cloud

could also be a message posted on a social media or a blog. Figure 4 illustrates the top 10 sources for both real and fake news articles. With respect to real news, we can notice that "Facebook posts", "Viral images" and "Blog posting" are the three main sources of fake news articles.



(a) Real News

(b) Fake News

Fig. 4: Top 10 News Sources

3.2 Preparing the Tweets Collection

In this stage, we track the diffusion of the news articles on the Twitter microblog, by collecting tweets of users that discuss or interact with these events. Because we aim to build a multimedia dataset with images, we only keep news articles with images attached.

We start by constructing two sets of the most representative keywords from (1) the title and (2) the body text of each article. This way of proceeding is motivated, on the one hand, by the fact that the titles used by the journalists are not always significant and, on the other hand, by the intention to capture all the keywords of the news articles appearing either in the title or in the body of the text.

For the title, we develop a natural language processing pipeline to extract key words by removing special tokens and common stop words. In this step, we kept at most six key words to avoid over general queries. Given the relatively large size of the body text, we first use the TextRank extractive text summarization algorithm [10]. The advantage of using an extractive algorithm rather than an abstractive one is to keep the original keywords of the text and the general meaning. We then apply our natural language processing pipeline to clean up the text. Note that we keep the named entities in both sets of keywords, which plays an important role for the semantics of the search queries. For example, Figure 6 shows the keywords extracted from the title and the body text of a news article.

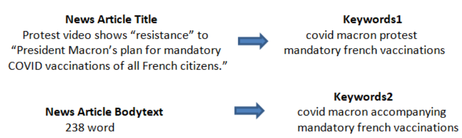


Fig. 5: An Example of Keyword Extraction

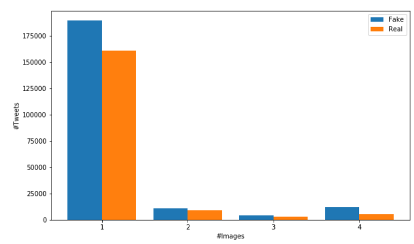


Fig. 6: Image Distribution

3.3 Tweets Collection

We have prepared two keyword sets that serve as queries to search for matching tweets. We next use the *Twitter Academic research API*[11] with the Full-archive search endpoint to collect the complete history of public tweets from March 2006 (Twitter creation) until September 07, 2021 (for the current version of the dataset). With a 1024 character limit on the query rules, this endpoint also supports a more advanced query language (Boolean) to help yield more precise, complete, and unbiased filtered results. Thanks to these features, we were able to create a single search query that combines the two keyword lists that were prepared earlier for each news article.

The Full-archive search endpoint returns matching tweets with complete information such as their IDs, text, date and times of being created. It also enables us to pull valuable details about users (e.g., their IDs) and their social activity

[11] https://developer.twitter.com/en/products/twitter-api/academic-research/product-details

(e.g., the number of retweets/followers/friends) and media content (e.g., images, places, and poll metadata) using the new expansion parameters. We also consider that all the tweets that discuss a particular news article will bear the truth value (i.e., the label of the article) because it contributes to the diffusion of a rumor (true or fake), even if the tweet denies or remains skeptical regarding the veracity of the rumor. The statistical details of DAT@Z21 are shown in Table 2.

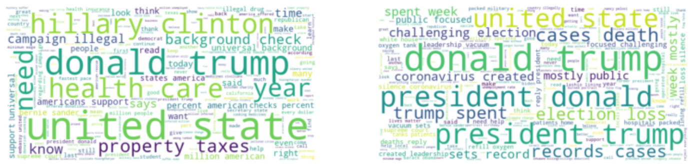
Table 2: Dataset Statistics

Characteristics	Real	Fake	Overall
All News articles	5,671	7,213	12,884
#With images	1,765	2,391	4,156
All Tweets	1,209,144	1,655,386	2,,864,530
#With images	179,153	216,472	395,625
#Images	211,447	271,419	482,866
#Tweets as retweets	4,734	5,502	10,236
#Tweets as replies	37,593	55,937	93,530
#Tweets as likes	2,803	4,453	7,256
#Tweets with location	3,823	4,497	8,320
#Users	98,967	115,669	194,692
Average #Followers	107,786	286,101	205,353
Average #Followees	2,929	3,153	3,051

Unlike most existing datasets that are focused on only text content, we aim to build a multimedia dataset that includes images. We collect the original tweet texts and attached images. Thus, from the 2,864,530 collected tweets, we remove text-only tweets, duplicated tweets, and images to obtain 395,625 tweets with attached images. This represents a significant part ($\simeq$ 14%). In addition, among these tweets, almost 10% share more than one image, which shows that visual information is widely used on social media. Figure 6 shows the distribution of true and fake tweets with the number of included images.

For the textual content of our dataset, we analyze the topic distribution of fake and real tweets. From Figure 7, we can observe that the topic distribution of fake tweets is slightly different from that of the real tweets. Indeed, both distributions share most of the frequent words related to political events, such as "Donald", "trump", "Hillary", "clinton", and "election" or to the Covid-19 health crisis, such as "health", "care", "Coronavirus", and "death". The accordance of real and fake news topic distributions ensures that the fake news detection models trained on our dataset are not topic classifiers.

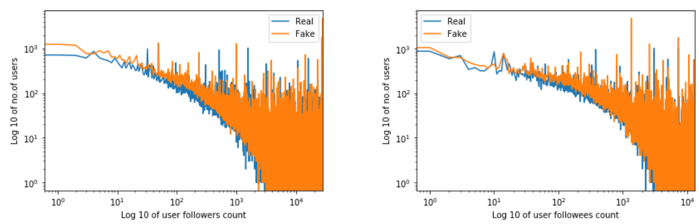
For users involved in spreading tweets, we find that the number of users spreading fake and true news is greater than the total number of distinct users in the dataset. This indicates that a user can engage in spreading both fake and real news. As for the users social interaction, the distribution of their followers and followees count are presented in Figure 8. We can observe that the followees and



(a) Real Tweets (b) Fake Tweets

Fig. 7: Topic Distribution of Fake and Real Tweets

followers count of users spreading fake and real tweets follows a distribution that resembles a power law, which is commonly observed is social network structures.



(a) Follower Count of Users (b) Followee Count of Users

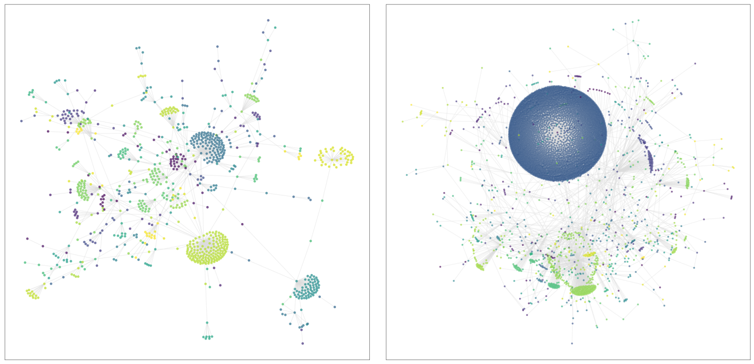
Fig. 8: Followers and Followees Count Distribution of Fake and Real Tweets

To understand how users interact with each other, we have created an interaction graph using the Python package Networkx [6] that represent three types of interactions between two users: retweets, replies, and quoted tweets. Figure 9 shows the largest connected component for our original graph for users spreading real and fake tweets. We note that both users display echo-chamberness. In addition, users posting fake tweets tend to create tightly knit groups characterized by relatively highly-dense ties.

Our dataset includes spatiotemporal information that depicts the temporal user engagement for news articles, which provides the necessary information to further investigate the utility of using spatiotemporal information to detect fake news. Figure 10 depicts the geolocation distribution of users engaging in posting tweets. Given that it is often rare for the location to be explicitly provided by the users in their profiles, we consider the locations information attached with tweets.

Note that to comply with Twitter's Terms of Service[12], the full contents of user social engagements and network are not able to be directly published.

[12] https://developer.twitter.com/en/developer-terms/agreement-and-policy



(a) Real Tweets (b) Fake Tweets

Fig. 9: Social Connections Graph for Real and Fake Tweets

Fig. 10: Spatial Distribution of Users posting Fake and Real Tweets

Instead, we only publicly release the collected tweet IDs, so that the tweets can be rehydrated for non-commercial research use, but we also provide the instructions for obtaining the tweets using the released IDs for user convenience. More details can be seen in https://git.msh-lse.fr/eric/dataz21. We also maintain and update annually the repository to ensure its usability.

4 Rumor Classification Using DAT@Z21

In this section, we conduct comparative experiments on the rumor classification task using the DAT@Z21 dataset. These experiments aim to show the usefulness of our dataset with different kinds of methods. We include both monomodal methods and multimodal state-of-the-art methods as baselines to predict the credibility of tweets. We specify these methods in Section 4.1. We then provide the experiments settings in Section 4.2. Finally, we present the performance analysis for these methods in Section 4.3.

4.1 Baselines

For monomodal methods, we consider the following algorithms:

NB (Naive Bayes) We use the doc2vec embedding model to represent texts and feed the representations to a Gaussian Naive Bayes algorithm.

LR (Logistic Regression) We use the word2vec embedding model to represent texts and feed the representations to the model.

RF (Random Forest) We use the Term Frequency-Inverse Document Frequency (TF-IDF) to represent texts and feed them to the model.

Text-CNN *[8]* This is a convolutional neural network for text classification, we feed the word embedding into a Convolution1D which will learn filters. We then add a max pooling layer.

For multimodal methods, we consider the following state of the art baselines:

SAFE *[22]* [13] This is a neural-network-based method that utilizes news multimodal information for fake news detection, where news representation is learned jointly by news textual and visual information along with their relationship (similarity). SAFE facilitates recognizing the news falseness in its text, images, and/or the "irrelevance" between text and images.

att-RNN *[7]* This is a deep model that employs LSTM and VGG-19 with attention mechanism to fuse textual, visual and social-context features of tweets. We set the hyper-parameters as in [7].

deepMONITOR *[1]* This is a multi-channel deep model where first a Long-term Recurrent Convolutional Network (LRCN) is used to capture and represent text semantics and sentiments through emotional lexicons. The VGG19 model is used to extract salient visual features from post images. Image features are then fused with the joint representations of text and sentiment to classify messages.

4.2 Experiment Settings

We randomly divided the dataset into training and testing subsets, with a ratio of 0.7:0.3. We use the scikit-learn library from Python to implement the traditional algorithms with default settings and we do not tune parameters. We use the Keras library from Python to implement multimodal methods. We use TF-IDF with 1000 features, word2vec and doc2vec embeddings with 200 vector size

[13] https://github.com/Jindi0/SAFE

to represent texts. Because the dataset has a slightly unbalanced distribution between false and true classes ($\simeq 1.17 : 1$), we use the following metrics to evaluate the performance of rumor classification algorithms: PR AUC (Area Under Precision-Recall Curve), ROC AUC(Area Under Receiver Operating Characteristic Curve), precision, recall, and F_1-score. Finally, we run each method three times and report the average score in Table 3.

Table 3: Classification Performance Using the DAT@Z21 Dataset

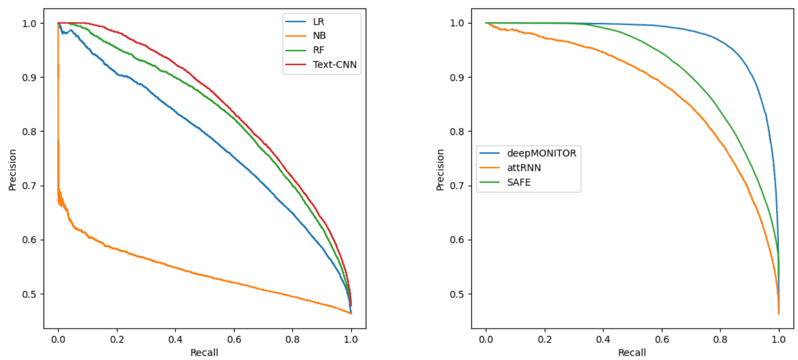
Type	Methods	PR AUC	ROC AUC	Precision	Recall	F_1
Monomodal methods	LR	0.778	0.799	0.729	0.777	0.752
	NB	0.496	0.561	0.593	0.480	0.531
	RF	0.843	0.846	0.754	0.783	0.764
	Text-CNN	0.861	0.866	0.759	0.774	0.771
Multimodal methods	att-RNN	0.880	0.871	0.805	0.753	0.777
	SAFE	0.937	0.935	0.851	0.857	0.854
	deepMONITOR	0.984	0.982	0.901	0.948	0.924

4.3 Experimental Results

From Table 3, we can see that among the four simple models, the Text-CNN is the best performing model and it achieved the best scores for four out of five comparison metrics, including PR AUC, ROC AUC, precision, and F_1-score. For multimodal techniques, deepMONITOR outperforms other models, with the best scores in all comparison metrics. Multimodal methods generally perform better than monomodal methods because they incorporate signals from other message modalities then the text content like visual, sentiment or social context information, which enables them to better capture contextual information.

Because the dataset is slightly unbalanced in favor of fake labels, the models tend to generate slightly more false positive cases. However, the PR AUC (which is an appropriate metric for unbalanced datasets), recall and F_1-score values are satisfactory. Figure 11 illustrates the PR curve for the monomodal and multimodal methods. Given that the proportion of true and fake information is likely to be even more unbalanced in real world datasets, practical detection solutions need to handle this type of imbalance more effectively.

With these experiments, we show that the DAT@Z21 dataset allows to explore various methods dealing with different types of features, by comparing their efficiency. According to the obtained results, we see that the visual content is very important since these methods obtain better results. Thus, a dataset containing all these features was required.



(a) Monomodal Methods (b) Multimodal Methods

Fig. 11: PR curve for Monomodal and Multimodal Methods

5 Conclusion and Perspectives

In this work, we have constructed a comprehensive multimodal dataset DAT@Z21 for rumor classification, which contains textual, visual, spatiotemporal, user engagement, and social platform posts. We have described how we collected the dataset. To show the distinctive features between rumors and facts, we have provided a brief exploratory data analysis of news articles and the corresponding messages from the Twitter social platform. In addition , we have demonstrated the usefulness and the relevance of our dataset through a rumor classification task over several monomodal and multimodal state-of-the-art methods using the data of DAT@Z21.

The DAT@Z21 dataset can be extended by including (1) news articles and tweets from other languages than English, and (2) other reliable news sources(e.g., other fact-checking or organizations and platforms that assess the reliability and bias of news sources) to create a large, centralized set of ground truth labels. We hope that researchers will find DAT@Z21 useful for their own research and that together we contribute to combat rumors.

References

1. Azri, A., Favre, C., Harbi, N., Darmont, J., Noûs, C.: Calling to cnn-lstm for rumor detection: A deep multi-channel model for message veracity classification in microblogs. In: ECML PKDD. pp. 497–513. Springer (2021)
2. Cao, J., Guo, J., Li, X., Jin, Z., Guo, H., Li, J.: Automatic rumor detection on microblogs: A survey. arXiv preprint arXiv:1807.03505 (2018)
3. Cui, L., Lee, D.: Coaid: Covid-19 healthcare misinformation dataset. arXiv preprint arXiv:2006.00885 (2020)

4. Dai, E., Sun, Y., Wang, S.: Ginger cannot cure cancer: Battling fake health news with a comprehensive data repository. In: Proceedings of the AAAI ICWSM. vol. 14, pp. 853–862 (2020)
5. Dong, E., Du, H., Gardner, L.: An interactive web-based dashboard to track covid-19 in real time. The Lancet infectious diseases **20**(5), 533–534 (2020)
6. Hagberg, A., Swart, P., S Chult, D.: Exploring network structure, dynamics, and function using networkx. Tech. rep., Los Alamos National Lab.(LANL), Los Alamos, NM (United States) (2008)
7. Jin, Z., Cao, J., Guo, H., Zhang, Y., Luo, J.: Multimodal fusion with recurrent neural networks for rumor detection on microblogs. In: ICM 2017. pp. 795–816. ACM (2017)
8. Kim, Y.: Convolutional neural networks for sentence classification. In: Proceedings of EMNLP. pp. 1746–1751. ACL, Doha, Qatar (2014)
9. Latif, S., Usman, M., Manzoor, S., Iqbal, W., Qadir, J., Tyson, G., Castro, I., Razi, A., Boulos, M.N.K., Weller, A., et al.: Leveraging data science to combat covid-19: A comprehensive review. IEEE Transactions on AI **1**(1), 85–103 (2020)
10. Mihalcea, R., Tarau, P.: Textrank: Bringing order into text. In: Proceedings of the 2004 conference on EMNLP. pp. 404–411 (2004)
11. Mitra, T., Gilbert, E.: Credbank: A large-scale social media corpus with associated credibility annotations. In: Ninth AAAI ICWSM (2015)
12. Nørregaard, J., Horne, B.D., Adalı, S.: Nela-gt-2018: A large multi-labelled news dataset for the study of misinformation in news articles. In: Proceedings of the AAAI ICWSM. vol. 13, pp. 630–638 (2019)
13. Pathak, A.R., Mahajan, A., Singh, K., Patil, A., Nair, A.: Analysis of techniques for rumor detection in social media. Procedia Computer Science **167**, 2286–2296 (2020)
14. Salem, F.K.A., Al Feel, R., Elbassuoni, S., Jaber, M., Farah, M.: Fa-kes: A fake news dataset around the syrian war. In: Proceedings of the AAAI ICWSM. vol. 13, pp. 573–582 (2019)
15. Shahi, G.K., Nandini, D.: Fakecovid–a multilingual cross-domain fact check news dataset for covid-19. arXiv preprint arXiv:2006.11343 (2020)
16. Shu, K., Mahudeswaran, D., Wang, S., Lee, D., Liu, H.: Fakenewsnet: A data repository with news content, social context and spatialtemporal information for studying fake news on social media. arXiv preprint arXiv:1809.01286 (2018)
17. Shu, K., Sliva, A., Wang, S., Tang, J., Liu, H.: Fake news detection on social media: A data mining perspective. ACM SIGKDD Explorations Newsletter **19**(1), 22–36 (2017)
18. Shuja, J., Alanazi, E., Alasmary, W., Alashaikh, A.: Covid-19 open source data sets: a comprehensive survey. Applied Intelligence pp. 1–30 (2020)
19. Tacchini, E., Ballarin, G., Della Vedova, M.L., Moret, S., de Alfaro, L.: Some like it hoax: Automated fake news detection in social networks. arXiv preprint arXiv:1704.07506 (2017)
20. Wang, W.Y.: " liar, liar pants on fire": A new benchmark dataset for fake news detection. arXiv preprint arXiv:1705.00648 (2017)
21. Zhou, X., Mulay, A., Ferrara, E., Zafarani, R.: Recovery: A multimodal repository for covid-19 news credibility research. In: Proceedings of the 29th ACM CIKM. pp. 3205–3212 (2020)
22. Zhou, X., Wu, J., Zafarani, R.: Safe: Similarity-aware multi-modal fake news detection. In: PAKDD conference. pp. 354–367. Springer (2020)